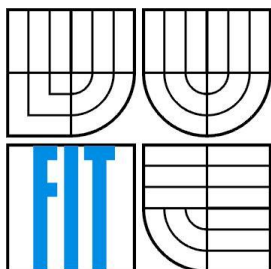


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

EDITOR ID3 TAGŮ S VYUŽITÍM METOD ZPRACOVÁNÍ OBRAZU

ID3 TAG EDITOR EMPLOYING IMAGE PROCESSING METHODS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

MAREK SEDLÁK

VEDOUCÍ PRÁCE

SUPERVISOR

ING. LUKÁŠ POLOK

BRNO

2010

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav počítačové grafiky a multimédií

Akademický rok 2009/2010

Zadání bakalářské práce

Řešitel: **Sedlák Marek**

Obor: Informační technologie

Téma: **Editor ID3 tagů s využitím metod zpracování obrazu**
ID3 Tag Editor Employing Image Processing Methods

Kategorie: Počítačová grafika

Pokyny:

1. Prostudujte základy zpracování obrazu, detekce objektů v obraze a OCR
2. Vytvořte přehled a hodnocení současných zdrojů informací o hudebních CD, zejména pak přebalů (cover art) a možnosti jejich využití
3. Navrhněte uživatelské rozhraní pro editor ID3 tagů s ohledem na automatizované hledání a korekce tagů
4. Implementujte navržený editor, podporující alespoň dva používané hudební formáty (ztrátový a bezztrátový)
5. Zhodnoťte dosažené výsledky a diskutujte možnosti budoucího vývoje.
6. Vytvořte stručný plakát prezentující vaši práci, její cíle a výsledky.

Literatura:

- dle pokynů vedoucího

Při obhajobě semestrální části projektu je požadováno:

- Body 1-3.

Podrobné závazné pokyny pro vypracování bakalářské práce naleznete na adrese

<http://www.fit.vutbr.cz/info/szz/>

Technická zpráva bakalářské práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap (20 až 30% celkového rozsahu technické zprávy).

Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Polok Lukáš, Ing.**, UPGM FIT VUT

Datum zadání: 1. listopadu 2009

Datum odevzdání: 19. května 2010

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta informačních technologií
Ústav počítačové grafiky a multimédií
612 66 Brno, Božetěchova 2
L.S.



doc. Dr. Ing. Jan Černocký
vedoucí ústavu

Abstrakt

Tato práce se zabývá spojením úpravy tagů v hudebních souborech a využití metod rozpoznávání obrazu pro snazší práci s obaly alb.

V úvodu je popsána distribuce hudby, různé podoby uchovávání a přehrávání hudby. Je zde zmíněn i problém legality stahování a distribuce hudby a výměnných sítí.

Dále se text věnuje metodám použitým při transformacích obrazu použitých v aplikaci.

Součástí práce je i příprava, implementace a demonstrace aplikace.

V závěru práce je diskutováno možné použití nebo budoucí rozšíření.

Abstract

This thesis discusses the co-implementation of music tags editing and image processing methods for comfortable maintenance of album arts.

The introduction deals with music distribution, different ways of storing and playing the music. Legality and peer-to-peer issue is also mentioned.

Furthermore, the thesis deals with image transforming methods used in the application.

A part of the work is the preparation, implementation and development of the application.

The conclusion discusses future usage or extensions.

Klíčová slova

ID3, MP3, Wave, FreeDB, Sobel, Hough, zpracování obrazu, hudba

Keywords

ID3, MP3, Wave, FreeDB, Sobel, Hough, image processing, music

Citace

SEDLÁK, Marek: *Editor ID3 tagů s využitím metod zpracování obrazu*, Brno, FIT VUT v Brně, 2010

Editor ID3 tagů s využitím metod zpracování obrazu

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Lukáše Poloka. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Marek Sedlák

17. 05. 2010

Poděkování

Rád bych poděkoval Ing. Lukášovi Polokovi, který byl vždy nápomocný a ochotný konzultovat a odborně mi vždy poradil. Také bych se rád poděkoval svým rodičům, kteří mi dali možnost studovat na VUT.

© Marek Sedlák, 2010

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

Obsah.....	1
1 Úvod.....	3
2 Distribúcia a uchovávanie hudby.....	4
2.1 Fyzické prenosné médiá	4
2.1.1 Audiokazeta	4
2.1.2 CD.....	4
2.1.3 MiniDisc	5
2.1.4 DVD.....	6
2.1.5 Blu-ray	6
2.1.6 Ochrana proti kopírovaniu	6
2.1.7 Galéria fyzických prenosných médií	7
2.2 Elektronické médiá	7
2.2.1 Internetové obchody	7
2.2.2 Peer-to-peer siete a autorské práva	8
2.3 Uchovávanie hudby	10
2.3.1 Úložné médiá	10
2.4 Formáty.....	11
2.4.1 Bezstratové	11
2.4.2 Stratové	12
Prehrávanie.....	14
2.5 Softvérové prehrávače	14
2.6 Databázy albumov	15
2.6.1 Freedb	15
2.6.2 Gracenote.....	15
2.6.3 Cdcovers.cc.....	15
3 Popis teórie.....	16
3.1 Tagy	16
3.1.1 ID3	16
3.1.2 Ostatné	18
3.2 Obraz ako signál	18
3.2.1 Kompresia obrazov	18
3.3 Farebné modely	19
3.3.1 Grayscale	19
3.3.2 RGB	19
3.3.3 YUV.....	20
3.3.4 $Y_{P_R}P_R$ a $Y_{C_R}C_R$	20
3.4 Detekcia hrán	20
3.4.1 Konvolúcia.....	20
3.4.2 Sobelov operátor.....	21
3.4.3 Prewittov operátor	23
3.4.4 Cannyho hranový detektor.....	23
3.5 Houghova transformácia.....	24
4 Požiadavky	27
5 Návrh aplikácie	27

5.1	Použité knižnice.....	27
5.1.1	LAME.....	27
5.1.2	TagLib#.....	29
5.1.3	Interop (COM objekty)	29
5.1.4	Ripper, Yeti.mmedia a Yeti.mp3	30
5.2	Grafické prvky	30
5.3	Vývojový nástroj	30
5.3.1	Technológia .NET.....	30
5.4	Implementácia.....	31
5.4.1	mainForm.cs	31
5.4.2	imageRecognition.cs.....	36
5.5	Test aplikácie	38
6	Záver	39

1 Úvod

Hudba sa vo všeobecnosti definuje ako špecifická ľudská aktivita, ktorá sa pomocou v priestore a čase charakteristicky zoskupených tónov a zvukov a na základe spoločenských skúseností usiluje o (hlavne estetickú) komunikáciu. Hudba je konkrétnejšie druh umenia, ktorého výrazovými prostriedkami sú tóny a základnými komponentmi harmónia, melódia, rytmus a farba. Podľa rímskeho filozofa Cassiodora je hudba matematická veda. [1]

Je to pojem, ktorý nie je cudzí snáď nikomu. Mnohí ľudia žijú hudbou, iní ju počúvajú na pozadí, prípadne ich stimuluje k vyšším výkonom alebo upokojuje. Každý z nich ju však musí mať k dispozícii, aby ju mohol počúvať. V dnešnej dobe je pojem „počítačová gramotnosť“ už niekde v pozadí, s počítačom totiž pracuje drvivá väčšina ľudí (hlavne) produktívneho veku. Každý má možnosť si takto uchovať okrem svojich textových prác, tabuliek, filmov takisto aj práve spomínanú hudbu. Nemalou mierou k tomu prispievajú aj prenosné prehrávače hudby, prípadne mobilné telefóny. Spríjemňujú ľuďom život počas cestovania, čakania a iných činnostiach. Hudba je všade, no prenosné médiá, prípadne prehrávače majú vždy určité obmedzenia. Kapacitné obmedzenie rieši uchovávanie svojich kolekcií na osobnom počítači, prípadne notebooku, kde je kapacita spravidla neporovnateľná.

Na trhu existuje množstvo správcov hudobných kolekcií, kvalitných, aj viac kvalitných, ktoré ponúkajú určitú automatizáciu. To zahŕňa napríklad automatické zistenie interpreta, názvov stôp, stiahnutie obalu albumu. Mnohokrát sú však pre bežného užívateľa máťúce a komplikované. Cieľom tohto projektu je zjednodušenie aplikácie tohto typu, pričom miera automatizácie ostane zachovaná na vyššej úrovni. Text sa najskôr venuje hudobnej časti, neskôr problematike rozpoznávania obrazu.

V 2. kapitole je popísaná distribúcia, ukladanie a uchovávanie hudby. Popisuje rôzne médiá a formy archivácie hudby do súborov. V 4. kapitole prevažne rozoberám obraz, jeho farebné modely a transformačné metódy. Spomenuté sú aj spôsoby ukladania informácií o albumoch v hudobných súboroch. Aplikácii sa venuje kapitola 5, kde sú spomenuté použité prvky a technológie. Je tu rozobraný návrh, implementácia a popísané testovanie aplikácie. Záverečná kapitola popisuje možné použitie a budúce rozšírenie.

2 Distribúcia a uchovávanie hudby

V dnešnej dobe je produkováaná hudba dostupná v elektronickej forme na rôznych portáloch a elektronických obchodoch, kam ju hudobné nakladateľstvá distribuuujú. Stále je však táto tvorba dostupná fyzicky na klasických CD nosičoch, prípadne DVD, Blu-ray a iných typoch diskov. V tejto kapitole sa pozrieme na možné postupy, ako a kde získať hudbu. Rozoberieme typy médií, porovnáme ich parametre.

2.1 Fyzické prenosné médiá

Táto podkapitola chronologicky prechádza jednotlivé typy.

2.1.1 Audiokazeta

Audiokazeta (Obr. 2.1) alebo kompaktná kazeta je magnetický dátový nosič vo forme magnetickej pásky úplne ukrytej v normalizovanom puzdre. Na pásku je možné analógovo zaznamenať zvuk v tzv. kazetových magnetofónoch alebo kazetových diktafónoch. Páska patrila medzi 70-tymi a 90-tymi rokmi k najpoužívanejším nosičom pre amatérsky záznam zvuku. Doteraz sa niektoré zvukové nahrávky vydávajú aj na kompaktnej kazete. [2]

2.1.2 CD

Kompaktný disk (skrátene CD) (Obr. 2.1) je optický dátový nosič. Pôvodne bol určený len na uloženie hudby v digitálnej forme. Neskôr tento nosič umožnil aj uloženie iných dát (CD-ROM), videozáznamu (VCD, SVCD) alebo fotografií (Photo CD). Neskôr boli vyvinuté aj prepisovateľné disky CD-RW. Audio CD (skratka CD-DA, CD - Digital Audio) je v súčasnosti najrozšírenejší spôsob distribúcie hudby na fyzickom médiu. Sú vyrobené z polykarbonátového plastu a sú hrubé 1,2 mm. Disky sú pokryté veľmi tenkou hliníkovou vrstvou (pôvodne sa používalo zlato, s ktorým sa môžeme stretnúť aj dnes, vďaka trvácnosti údajov uložených na vrstvu zo zlata). Hliníková vrstva je chránená lakovým filmom. Na lakový film sa môže vytlačiť potlač disku. Kompaktné disky sa vyrábajú v rôznych veľkostiach (priemeroch) a dokonca aj tvaroch (tzv. eng. shape disc, v preklade tvarovaný disk), najčastejšie však majú disky priemer 120 mm, ktorých kapacita býva 80 minút pre záznam zvuku alebo 700 MB pre uloženie údajov.

Niektoré ponúkané prázdne disky uvádzajú kapacitu 90 alebo 99 minút. Špecifikácia pre zapisovateľné kompaktné disky uvádza maximálnu kapacitu 74 minút a kapacita do 80 minút je maximálna možná odchýlka. Preto všetky kompaktné disky s kapacitou nad 80 minút porušujú túto špecifikáciu a viacero prehrávačov alebo rekordérov môže mať problém disky s kapacitou 90 a viac minút používať. Red Book (Červená kniha) je štandard pre audio CD disky (tj. zvukové CD). Jej

názov vyplýva z prebalu knihy zo sady tzv. Rainbow Books (Dúhových kníh), ktoré definujú používané formáty kompaktných CD diskov.

2.1.2.1 Špecifikácia a technické detaily Audio-CD podľa Red Book

- Maximálny hrací čas je 79.8 minúty
- Minimálna dĺžka stopy sú 4 sekundy (vrátane 2-sekundovej pauzy)
- Maximálny počet stôp je 99
- Maximálny počet indexových bodov je 99 bez časového limitu
- Medzinárodný štandardný nahrávací kód (ISRC) by mal byť zahrnutý v disku

Red Book špecifikuje fyzické parametre a vlastnosti CD, optické parametre, výnimky a chybovosť, regulačný systém (EFM) a opráv chýb (CIRC), a kanály subkódu a grafiku. Tiež určuje formu digitálneho kódovania: 2-kanálové signed 16-bit lineárne PCM vzorky na 44,100 Hz. Táto vzorkovacia frekvencia je adaptovaná z ekvivalentu počas nahrávania digitálneho audia na PAL videokazetu s PCM adaptérom (predchádzajúci spôsob ukladania digitálneho audia).

Prenosová rýchlosť je 1411,2 kbit/s; vzorec:

$$2 \text{ kanály} \times 44,100 \text{ vzoriek za sekundu na kanál} \times 16 \text{ bitov na vzorku} = 1,411,200 \text{ bitov za sekundu} = 1,411.2 \text{ kbit/s.}$$

Každá vzorka je „signed“ 16-bitové číslo, t.j. v rozsahu -32768 to až 32767. Na disku sú dáta uložené v sektoroch po 2352 bajtoch, čítané rýchlosťou 75 sektorov za sekundu. K tomu treba prirátat' EFM, CIRC, L2 ECC, tieto údaje však zvyčajne nie sú viditeľné pre aplikácie. Pre porovnanie, prenosová rýchlosť dátového CD "1x" je definovaná ako 2048 bajtov za sektor \times 75 sektorov za sekundu = 150 KB/s, alebo tiež približne 9,2 milióna bajtov za minútu. [3] [7]

2.1.3 MiniDisc

Minidisc (MD) (Obr. 2.1) je magnetooptické médium uvedené na trh firmou Sony v apríli 1991 a to primárne pre nahrávanie a distribúciu hudby. V roku 1993 bol na trh uvedený aj minidisc s možnosťou ukladať dáta. Kapacita média je podobne ako u CD 74 alebo 80 minút hudby alebo zvuku. Pojem Minidisc sa tiež v skratke označujú rekordéry alebo prehrávače, ktoré s týmto médiom pracujú. Tie používajú k ukladaniu zvukovej nahrávky stratovú kompresiu (podobne ako MP3) s názvom ATRAC (Adaptive TRansform Acoustic Coding). Od uvedenia na trh sa objavilo niekoľko variantov MD prehrávačov – okrem štandardného MD aj MDLP, NetMD a Hi-MD. [4] Jeho hlavná výhoda oproti CD spočívala v menšej veľkosti a väčšej odolnosti. Avšak tento formát sa na trhu neujal.

2.1.4 DVD

DVD (Obr. 2.1) je formát digitálneho optického dátového nosiča, ktorý môže obsahovať filmy vo vysokej obrazovej a zvukovej kvalite alebo rozličné iné údaje. Bol vyvinutý a vynájdený spoločnosťami Sony a Philips v roku 1995. Používa sa hlavne na ukladanie digitálneho videa a rôznych digitálnych dát. Disky DVD majú rovnaké rozmery ako kompaktné disky CD, ale dokážu uložiť až šesť krát viac dát. Disk DVD sa na pohľad podobá kompaktnému disku.

DVD-Audio (skratka DVD-A) je digitálny formát, určený na ukladanie vysoko kvalitného zvukového záznamu na médiách typu DVD. DVD-Audio neobsahuje video, nemalo by sa teda zamieňať s Video-DVD, ktoré obsahuje hudobné videá. Prvé disky v tomto formáte vstúpili na trh v roku 2000. Dáta na týchto diskoch sú vo formáte *.war, rovnako ako na Super Audio CD (SACD), čo je iný formát na reprodukciu vysoko kvalitného zvuku. [5]

2.1.5 Blu-ray

Blu-ray Disk (oficiálna skratka BD, tiež známe ako BR alebo Blu-ray) (Obr. 2.1) je optické diskové pamäťové médium určené ako náhrada štandardu DVD. Používa sa hlavne pre ukladanie videa s vysokým rozlíšením, hier, a ďalších údajov. Jednovrstvové médium má kapacitu 25 GB, a dvojvrstvové 50 GB. Mimo štandardu však už existujú aj 200 GB disky (8 vrstiev), na ktoré ale treba špeciálny firmware alebo špeciálne vybavenie. Disky s takisto neštandardnou veľkosťou 100 GB však zvládne zariadenie aj bez špeciálneho firmwaru. Čo sa týka audia, Blu-Ray prehrávače musia podporovať Dolby Digital (AC-3), DTS, a lineárne PCM. Voliteľne môžu podporovať Dolby Digital Plus a DTS-HD Audio vo vysokom rozlíšení, tak ako aj bezstratové formáty Dolby TrueHD a DTS-HD Master Audio. [6]

2.1.6 Ochrana proti kopírovaniu

Väčšina týchto ochrán je zameraná na počítače s CD-ROM, tie spravidla musia podporovať tzv. Mixed CD formát a/alebo musia vedieť čítať Multisession-disky. V prípade Audio CD je ochrana mimo štandardu popísanom v Red Book. [7] Tieto ochrany sú spravidla prelomiteľné. Väčšina ochrán je tvorená pre Microsoft Windows systémy. Existuje softvér, ktorý buď ochranu prelomí real-time, alebo vytvorí obraz disku, a ten následne načíta do virtuálnej mechaniky už bez ochrany. Pre systém už je ďalej tento disk de facto nechránený.

2.1.7 Galéria fyzických prenosných médií



Obrázok 2.1: Ukážka prenosných médií
(zľava: Audiokazeta, CD, MD, DVD, BD)

2.2 Elektronické médiá

2.2.1 Internetové obchody

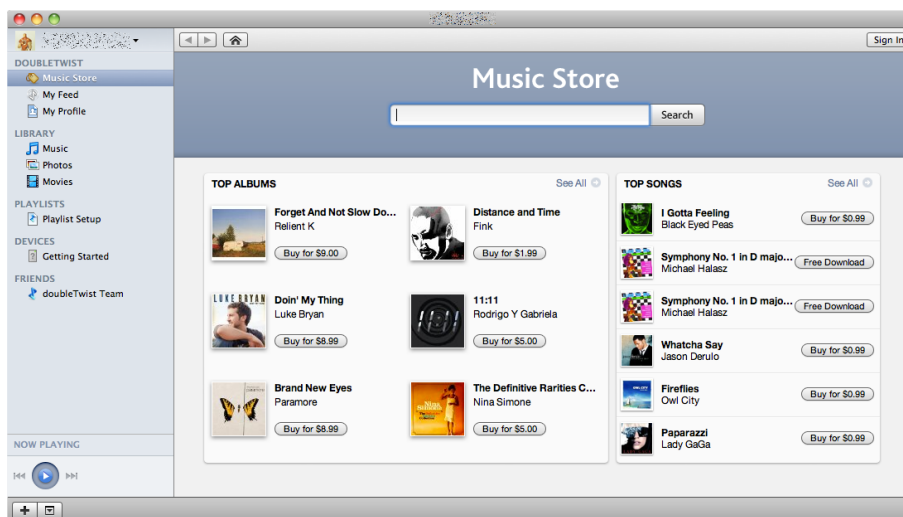
Hudba ponúkaná na internete je vo väčšine prípadov lacnejšia, než kúpená v obchode. Je to logické, keďže pri tomto druhu distribúcie si výrobca môže ako náklady odrátať plastový obal, potlač a zviazanie prebalu a potlač CD. Najznámejšie portály s touto službou sú Amazon a iTunes.

2.2.1.1 Amazon

Amazon.com je internetový obchod patriaci americkej spoločnosti Amazon.com, Inc. v štáte Washington. Ponuka je široká, užívateľské prostredie je rozdelené na hornú časť, kam zadávame hľadaný výraz, ľavú časť, ktorá obsahuje filter a pravú – hlavnú – časť, kde môže užívateľ vidieť buď detaily o albume alebo zoznam nájdených albumov. Hlavná časť obsahuje obrázok prebalu, zoznam skladieb s možnosťou kúpy po jednej, dĺžku trvania skladieb, cenu a iné detaily o albume.

2.2.1.2 iTunes Store

iTunes Store (Obr. 2.2) je on-line obchod s digitálnymi médiami, založený na vlastnom softvérovom riešení, prevádzkovaný spoločnosťou Apple. Dňa 24. februára 2010 sa v obchode predala 10 000 000 000-ta skladba na stiahnutie. iTunes predstavuje 70% celosvetového on-line predaja digitálnej hudby.[8] Do 6. januára 2009 sa väčšina hudby vzťahovala DRM (správa digitálnych práv - rôzne obmedzenia prehrávania týkajúce sa autorských práv). Vtedy Apple oznámil, že DRM bola vymazaná z 80% kompletného hudobného katalógu a prešiel na „DRM-free“ politiku. Takáto hudba je poskytovaná v rámci služby iTunes Plus.[9]



Obrázok 2.2: Aplikácia iTunes [10]

2.2.2 Peer-to-peer siete a autorské práva

Peer-to-peer, P2P alebo klient-klient je označenie architektúry počítačových sietí, v ktorej spolu komunikujú priamo jednotliví klienti (užívatelia). [11] V súvislosti s hudbou, filmami a softvérom sa tento pojem spomína pri výmenných sieťach. Užívatelia si takto jednoducho zdieľajú navzájom rôzny obsah. V Českej republike je sťahovanie hudby legálne za účelom vytvorenia vlastnej kópie a pre vlastné užitie. Nelegálne je poskytovanie, inak povedané šírenie obsahu chráneného autorskými právami tretej osobe. [11]

2.2.2.1 BitTorrent

BitTorrent je peer-to-peer internetový protokol, ktorý poskytuje rýchly, jednoduchý a efektívny prenos veľkých súborov medzi ľubovoľným počtom ľudí. Základný klient sa volá rovnako – BitTorrent. Najpoužívanjšie sú μ Torrent, Vuze (predtým Azureus), BitSpirit a iné. V princípe je sťahovanie obsahu chráneného autorskými právami nelegálne. Protokol BitTorrent totiž v rovnakom momente, v ktorom klient sťahuje dáta, tieto dáta dáva k dispozícii pre ostatných, a tým porušuje autorské právo.

Sťahovaný obsah je definovaný v *.torrent súboroch, ktoré sú k dispozícii na rôznych portáloch, tzv. *trackeroch*. Najznámejšie servery: The Pirate Bay, Mininova. Viacero takýchto stránok bolo pod žalobami nútené svoje pôsobenie zrušiť (TorrentSpy.com), iné odolávajú (The Pirate Bay) napriek početným žalobám od spoločností chrániacich autorské práva. Využívajú rozdielnu interpretáciu zákonov týkajúcich sa autorských práv v rozdielnych štátoch.

2.2.2.2 **Direct Connect**

Direct Connect je protokol pre peer-to-peer zdieľanie súborov. Základný a najrozšírenejší klient pre Windows je DC++. Kvalitný je český program StrongDC, ktorý podporuje aj 64-bitové inštrukcie a segmentované sťahovanie.

Princíp tejto siete je mať nazdieľaný obsah do určitej veľkosti (typicky 5 GB, 10 GB, 20 GB, 30 GB alebo 50 GB). Server sa v terminológii tohto protokolu nazýva *hub*. Ten si určuje pravidlá, kedy sa klient môže pripojiť. Často je to práve nazdieľané množstvo dát. Ak má používateľ nazdieľaný obsah chránený autorskými právami, v momente pripojenia sa stáva osobou, ktorá tento obsah šíri, a to je protizákonné.

V prípade, kedy nezdíeľame nič, alebo zdíeľame len obsah, ktorý môže byť šírený, môže byť pri vhodnom serveri využitie takéhoto klienta ľahký a legálny (aj keď nemorálny) spôsob, ako si zaobstarat' kolekcie hudby.

2.2.2.3 **Ostatné siete [12]**

- Napster - rozšírené hlavne v 90. rokoch
- KaZaa (protokol FastTrack) – rozšírené hlavne v 90. rokoch
- eMule (protokol eDonkey, Kademia)
- Shareaza (Gnutella, Gnutella2)
- BearShare

2.3 Uchovávanie hudby

V tejto podkapitole sú popísané možnosti uchovávania a archivovania dát.

2.3.1 Úložné médiá

2.3.1.1 Prenosné médiá

Optické médiá

Optické prenosné médiá sú podrobnejšie popísané v 2. kapitole. Komplikovanejší je prístup na takéto médiá. Čítanie a zápis v rovnakej chvíli nie je možný. Na optický disk sa typicky zapíše (tzv. *napáli*) obsah, ktorý sa dlhodobo nebude meniť. Potom je disk pripravený na čítanie dát. Tento problém čiastočne rieši technológia Multisession, ktorá dovoľuje dopaľovanie/prepaľovanie, nie je však kompatibilná so staršími operačnými systémami.

Prenosné prístroje

Dáta, a teda aj hudba, môže byť uložená aj priamo v prístroji v internej alebo v externej pamäti. Interná pamäť je typicky typu flash. K takýmto prístrojom sa radia napríklad:

- Prenosný MP3 prehrávač (neznamená, že prehráva len formát mp3; zaužívaný názov)
- Mobilný telefón
- Rôzne prenosné herné konzoly (Sony PSP, Nintendo DS, a iné)

Externú pamäť pri týchto prístrojoch spravidla predstavujú pamäťové karty.

Pamäťové karty

Pamäťová karta je elektronické zariadenie na ukladanie dát. Používa sa v digitálnych fotoaparátach, PDA, notebookoch, mobilných telefónoch, prehrávačoch, video hrách, digitálnych záznamníkoch a iných elektronických zariadeniach. Je založená na pamäti typu flash EEPROM. Je to malé, kompaktné zariadenie s relatívne vysokou kapacitou (oproti ostatným pamäťovým médiám), odolné pri zaobchádzaní (odolné voči magnetickým a elektrickým poliam). Karty boli navrhnuté ako náhrada pevného disku pre zariadenia, kde tento nie je možné z rozmerových dôvodov použiť.



Obrázok 2.3: Sony Memory Stick Micro (M2) ; rozmery 15 × 12.5 × 1.2 mm; kapacita 64 MB ~ 32 GB

2.3.1.2 Neprenosné médiá

Pevné disky

Pevný disk (z angl. hard disk, HD alebo harddisk drive, HDD) je zariadenie, ktoré sa používa na uchovávanie dát v počítačoch. V domácnosti sú pevné disky najtypickejšie médium, kde uchováваме svoje dáta, samozrejme aj hudbu. Výhoda tohto média je v kapacite. Hranica dnešných bežných harddiskov je na úrovni 1TB. Po zapnutí počítača môžeme takto mať k dispozícii mesiace nepretržitej hudby. Pevný disk je základné úložné médium aj u notebookov, vyskytuje sa aj v mobilných telefónoch, prípadne v externej forme pripojiteľnej cez USB port, takže sa dá radiť aj medzi prenosné médiá.

Sieťové médiá

Server (z angl. to serve - slúžiť, poskytovať službu) je softvér, ktorý v rámci počítačovej siete zastáva v hierarchii nadradené miesto vzhľadom na funkcionality. Poskytuje svoje prostriedky a výkon klientom. Dnešné prehrávače, softvérové aj hardwarové, dokážu prehrávať hudbu zo vzdialených zdrojov, serverov. Existujú servery, ktoré sú jednosmerné (server-klient) fungujú ako rádio, tj. klient len prijíma dáta (hudbu) a nedokáže ju ovplyvniť. Takéto služby alebo protokoly sú napríklad Shoutcast, MMS server, RTSP server alebo UDP Multicast server. Sieťové médiá, kde si používateľ môže kopírovať hudbu a následne podľa želania prehrávať, môžu byť použité servery na protokole HTTP, FTP, SMB, NFS, Windows share, prípadne vo Windows 7 zabudovaný server media server.

2.4 Formáty

2.4.1 Bezstratové

2.4.1.1 Waveform Audio (*.wav)

WAV (alebo aj WAVE) je skratka a bežne používaná prípona pre Waveform audio formát. Tento zvukový formát vytvorili firmy IBM a Microsoft pre ukladanie zvuku na PC. Je to špeciálny variant obecného formátu RIFF. [13] Napriek tomu je možné ukladať do WAV súboru zvuk komprimované, napríklad pomocou GSM kompresie, ADPCM, μ -Law, A-Law či v MP3, väčšinou sa používa nekomprimovaný zvuk v pulznej kódovej modulácii. Rovnakým spôsobom je uložený zvuk na Audio CD, čo umožňuje ľahký prevod medzi týmito formátmi. [14]

2.4.1.2 Free Lossless Audio Codec (*.flac)

Free Lossless Audio Codec (FLAC) je slobodný súborový formát určený na uchovanie audio dát. Je to bezstratový kompresný formát, teda na rozdiel od stratových formátov ako MP3, AAC alebo

Vorbis nedochádza pri kompresii k strate informácií. Veľkosť súboru skomprimovaného týmto kodekom je asi 30-50% veľkosti pôvodného súboru. [15]

2.4.1.3 **Monkey's Audio (*.ape)**

Monkey's Audio je bezstratový audio kodek s príponou .ape, nemusíte zaň nič platiť a jeho softvér je určený pre Windows, na Linuxu je podporovaný len okrajovo. Má najlepší kompresný pomer, to je ale vykúpené tým, že je relatívne pomalý. Výsledná veľkosť súborov býva približne polovičná oproti originálu a päťnásobná oproti MP3 s dátovým tokom 192 kbit/s (MP3 je ale stratový kodek). [16]

2.4.2 **Stratové**

2.4.2.1 **MPEG Layer-3 (*.mp3)**

MP3 (MPEG-1 Layer 3) je formát stratovej kompresie zvukových súborov, založený na kompresnom algoritme MPEG (Motion Picture Experts Group). Bol vynájdený a štandardizovaný v roku 1991. Pri zachovaní vysokej kvality umožňuje zmenšiť veľkosť hudobných súborov v CD kvalite približne na desatinu až dvanástinu.

Formát MP3 sa stal obľúbeným pre uchovávanie a prehrávanie hudby na počítačoch, v súčasnosti sa už vyrábajú aj stolné a prenosné prehrávače tohto formátu. O vývoj formátu MP3 sa zaslúžil nemecký vedec Karlheinz Brandenburg, riaditeľ pobočky Fraunhoferovho ústavu pre mediálnu komunikáciu v Erlangene, a jeho vedecký tím v snahe vytvoriť vhodný dátovo menej náročný štandard vhodný pre použitie v digitálnom rozhlasе DAB. MP3 sa snaží odstrániť redundanciu zvukového signálu na základe psychoakustického modelu. Teda zo vstupného signálu sa odoberú informácie, ktoré človek nepočuje, alebo si ich neuvedomuje. Využíva sa princíp časového a frekvenčného maskovania. Kompresia zvuku podľa štandardu MPEG-1 obsahuje 3 vrstvy, ktoré sa líšia kvalitou a obtiažnosťou implementácie. [17]

2.4.2.2 **OGG Vorbis (*.ogg)**

Vorbis je slobodný a open source stratový audio kodek, ktorý sa mal stať náhradou formátu MP3. Jeho vývoj je zastrešený organizáciou Xiph.Org Foundation. Najčastejšie býva uložený v Ogg a v tomto spojení je často nazývaný Ogg Vorbis. Kodek Vorbis sa stal prvým Xiph kodekom pripraveným pre všeobecné použitie, a v istom čase v podstate jediným. Tento kodek bol v minulosti bohužiaľ často zamieňaný s kontajnerovým formátom OGG. Kodek Vorbis je pomenovaný podľa postavy zo Zemeplochy. V knihe Malí bohovia je jednou z dôležitých postáv Diakon Vorbis. [18]

2.4.2.3 **Advanced Audio Coding (*.aac)**

Advanced Audio Coding (skrátene AAC) je štandard pre stratovú kompresiu zvuku. Bol vyvinutý ako logický nasledovník formátu MP3 na stredných až vyšších prenosových rýchlostiach v rámci

štandardu MPEG4. Formát AAC nie je úplne jednotný a obsahuje v sebe niekoľko profilov, vylepšení apod. Taktiež existuje mnoho enkodérov (väčšinou proprietárnych), ktoré sa razantne líšia kvalitou. Odhliadnuc od toho, AAC je jeden z najpokročilejších kodekov a má veľmi dobré vyhliadky do budúcnosti. Jeho rôzne modifikácie ako AACplus apod. obsahujú veľmi pokročilé technológie ako LTP (Long Term Prediction) alebo Postprocessing, ktoré z nich na nízkych prenosových rýchlostiach robia najlepšie kodeky (často lepšie než Vorbis). [19]

2.4.2.4 **Windows Media Audio (*.wma)**

Windows Media Audio (WMA) je komprimovaný zvukový formát vyvinutý ako súčasť Windows Media bol pôvodne určený ako náhrada za MP3 (ktoré bolo patentované a Microsoft musí platiť za jeho začlenenie vo Windows). Dnes skôr súperí s AAC.

Vývoj WMA sa dá rozdeliť do 2 fáz, a to do vydania Windows Media Player 9 a po vydaní WMP9. Do verzie 8 vrátane to bol kodek, ktorý si po kvalitatívnej strane takmer nevyslúžil pozornosť a nebol veľmi konkurencieschopný, no od verzie 9 sa dostal na úroveň najvyspelejších kodekov. Aktuálna verzia je 9.1, ktorá okrem pôvodného stratového kodeku pridáva aj kodeky pre bezstratovú a multikanálovú kompresiu. WMA súbory sú takmer výlučne v kontajneri ASF. [20]

Prehrávanie

2.5 Softvérové prehrávače

Softvérových prehrávačov je veľké množstvo. Rozdeliť by sa dali podľa operačných systémov. V dnešnej dobe je každý skinovateľný a podporuje množstvo formátov, vizualizácie, napaľovanie hudby ako Audio CD a prehrávanie videa. Zhrniem niekoľko najznámejších a najpoužívanějších.

Winamp (Win32) - relatívne rozšírený prehrávač, freeware, má aj platenú verziu. Dokáže napaľovať CD, podporuje rozsiahle množstvo formátov. Kvalitná podpora streamovaných médií.

Windows Media Player (Win32) - zabudovaný v operačných systémoch Microsoft Windows. Používa dekodéry zabudované v systéme.

iTunes (MacOS/Win32) – spojený s iTunes Store – obchod s hudbou a softvérom. Spoločnosť Apple tento produkt distribuuje ako softvér na úpravu hudobných kolekcií v ich prenosných prehrávačoch iPod a i.

VideoLan (Win32/MacOS/Linux) – orientovaný pôvodne na video. Je to veľmi kvalitný open-source produkt, ktorý v súčasnosti vyvíjajú kontribútori. Výnimočnosťou tohto riešenia je fakt, že obsahuje streamovací server.

QuickTime Player (MacOS/Win32) – zabudovaný v operačných systémoch od spoločnosti Apple.

Amarok (Linux/MacOS) – graficky pekne spracovaný prehrávač, podporuje vyhľadávanie textu piesne a informácií o albume.

Last.fm (Web) – on-line riešenie, kde si používateľ môže vypočuť hudbu legálne a zadarmo. Tento projekt by sa dal zaradiť aj medzi sociálne siete. Registrovaní užívatelia môžu zdieľať, akú hudbu počúvajú. Existujú zásuvné moduly do hore uvedených prehrávačov, ktoré aktualizujú užívateľský profil podľa prehrávanej hudby.

2.6 Databázy albumov

2.6.1 Freedb

Freedb je databáza na vyhľadávanie informácií o CD cez internet. Klientská aplikácia musí (približne) vypočítať ID disku a nasledovne vyslať požiadavku na databázu. Výsledok sa zobrazí ako názov albumu, meno interpreta, zoznam skladieb a ďalšie informácie. [21] Server načúva na portoch 80 a 8080 na adrese freedb.freedb.org. Databáza je aktualizovaná niekoľko krát denne. Jej použitie je zdarma pre komerčné aj nekomerčné účely. Používa sa v hudobných prehrávačoch alebo aj CD ripperoch. Poskytuje aj webové vyhľadávanie.

2.6.2 Gracenote

Gracenote MusicID poskytuje rýchly, presný a spoľahlivý hudobný obsah pre ľudí na celom svete. Technológii spoločnosti Gracenote dôveruje množstvo vedúcich svetových elektronických a softvérových spoločností. Gracenote MusicID požíva viac-krokovú metódu rozpoznávania tak, aby umožnila identifikáciu, kategorizáciu a organizáciu digitálnej hudby. Nezáležiac na formáte zdroja, Gracenote dáva hudobným fanúšikom nástroje na správu a užívanie ich hudobných kolekcíí. [22] Je to komerčný projekt, ktorý navyše poskytuje aj texty skladieb, monitoruje celosvetové rebríčky prehrávania, je zabudovaný napríklad aj v telefónoch Sony Ericsson pod službou TrackID, ktorá zistí názov skladby napríklad aj z počúvaného rádia. Zo známeho softvéru ho používa napríklad Winamp a iTunes.

2.6.3 Cdcovers.cc

Najväčší svetový archív CD a DVD prebalov (Obr. 3.2) [23]. Je to webový portál, ktorý poskytuje kompletne prebaly CD, DVD a rôznych iných formátov. Obaly sú uploadované registrovanými užívateľmi a preto tento zdroj nemožno považovať za 100% spoľahlivý.

3 Popis teórie

3.1 Tagy

3.1.1 ID3

ID3 tag bol primárne vyvinutý pre počítačový hudobný formát MP3, možno ho však aj inde. Napríklad OGG formát môže použiť ID3 tag, ale narušilo by to jeho štruktúru. Podľa tzv. ID3 tagov dokážu hudobné prehrávače zobraziť názvy interpretov, piesní, albumov atd. ID3 tag nijak nesúvisí s názvom súboru MP3. Tento tag možno upraviť v niektorom programe vytvorenom pre tento zámer.

3.1.1.1 ID3v1

Po vzniku MP3 štandardu sa objavil problém s ukladáním údajov o súbore. Samostatný MP3 formát nemal žiadny spôsob, ako to dosiahnuť. V roku 1996 Eric Kemp navrhol pridať dodatočné dáta do audio súboru, čo by problém vyriešilo. Štandard, pomenovaný ID3v1, sa rýchlo stal štandardom pre ukladanie metadát v MP3.

Formát bol vydaný spoločnosťou Damaged cybernetics, undergroundovou skupinou, ktorá sa špecializovala crackovaním systémov herných konzol. Tento formát bol prvýkrát použitý v roku 1996 na lepšie indexovanie v ROM 8-bitového Nintenda. Eric a jeho spolupracovníci prirodzene preniesli tento tag do MP3 súborov. Tento formát tagu bol použitý pre rôzne, vted nie moc rozšírené, formáty.

ID3v1 Tag zaberá 128 bytov, začínajúc reťazcom TAG. Tag bol umiestnený na konci súboru pre zachovanie kompatibility so staršími prehrávačmi. Tento tag umožňuje 30 bytov pre názov, interpreta, album, a komentár, 4 bajty pre rok, a 1 bajt na identifikáciu žánru piesne z preddefinovaného zoznamu 80 hodnôt (prehrávač Winamp neskôr rozšíril tento zoznam na 148 hodnôt).

Vylepšený tag bol uvedený v roku 1997 Michaelom Mutschlerom. Ten tvrdil, že pole, kde bol uložený komentár, bolo príliš krátke. Rozhodol sa teda skrátiť ho o 2 bajty a tie použiť na číslo skladby. Takéto tagy sú označované ako ID3v1.1. Rozšírený tag (Extended tag) (Tab. 5.1) je dodatočný blok dát uložený pred ID3v1 tagom, ktorý rozširuje pole názvu, interpreta a albumu o 60 bajtov, žáner definuje ako voľný text (namiesto preddefinovaného indexu), pole rýchlosť zaberá 1 bajt. ďalej sa tu objavuje pole začiatok a koniec skladby vo formáte mmm:ss (napríklad kvôli použitiu "fading" efektu), môžu byť vynechané. [24]

Pole	Dĺžka	Popis
hlavička	4	Identifikačný reťazec "TAG+"
názov	60	dodatočných 60 znakov v poli názvu skladby
interpret	60	dodatočných 60 znakov v poli interpreta
album	60	dodatočných 60 znakov v poli albumu
rýchlosť	1	0= nenastavené; {1,2,3,4}={pomalá, stredná, rýchla, hardcore}
žáner	30	30 znakov pre pole žánru
začiatok skladby	6	začiatok skladby vo formáte mmm:ss
koniec skladby	6	koniec skladby vo formáte mmm:ss

Tabuľka 4.1: Štruktúra bloku rozšíreného tagu (227 bajtov) [24]

3.1.1.2 ID3v2

V roku 1998 bol vytvorený nový štandard s názvom ID3v2. Aj keď nesie názov ID3, s pôvodným ID3 štandardom nemá nič spoločné. ID3v2 tagy sú premenlivej veľkosti, a zvyčajne sa vyskytujú na začiatku súboru, na podporu streamovania médií. Skladajú sa z niekoľkých rámcov, z ktorých každý obsahuje časť metadát. Tie môžu byť zaberať až 16MB, zatiaľ čo celková veľkosť tagu je obmedzená na 256 MB. Kódovanie pomocou UTF-16. Tento tag začína hlavičkou v podobe reťazca „ID3“. [25]

Uvedením tohto formátu tagu bol ID3v1 tag označený ako zastaralý.

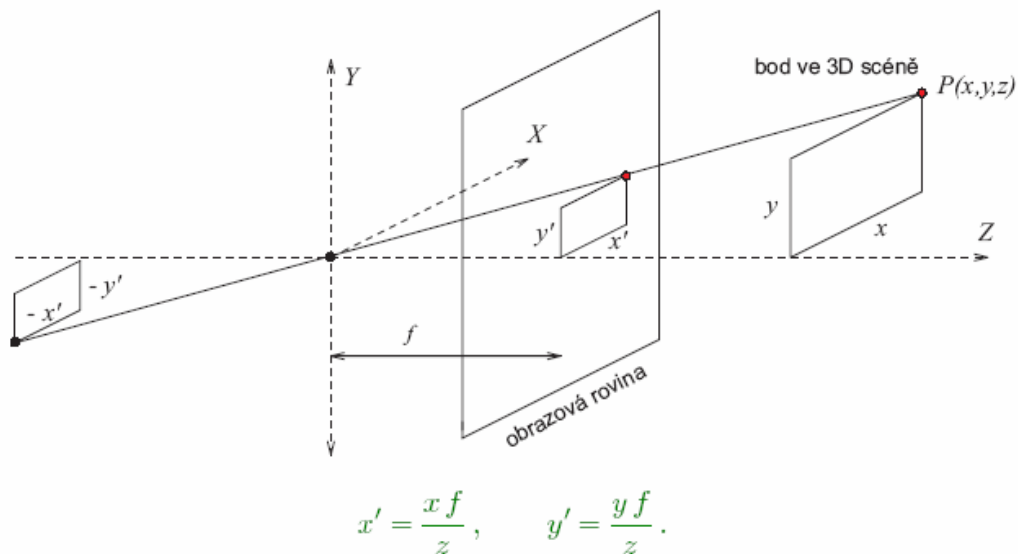
AENC	Audio kódovanie	TIT3	Úvodný kľúč
APIC	Priložený obrázok	TLAN	Jazyk(y)
COMM	Komentár	TLEN	Dĺžka
COMR	Komerčný rámec	TMED	Typ média
ENCR	Registrácia kódovania	TOAL	Názov originálneho albumu/média/relácie
EQUA	Ekvalizácia	TOFN	Pôvodný názov súboru
ETCO	Kódy načasovania udalostí	TOLY	Meno pôvodného autora textu
GEOB	Všeobecný zapuzdrený objekt	TOPE	Pôvodní umelci
GRID	Identifikácia registrácie skupiny	TORY	Pôvodný rok vydania
IPLS	Zoznam zahrnutých ľudí	TOWN	Vlastník súboru/licencia
LINK	Externá informácia	TPE1	Sólista
MCDI	Identifikátor hudobného CD	TPE2	Skupina
MLLT	Tabuľka lokalizácie formátu MPEG	TPE3	Dirigent
OWNE	Rámec týkajúci sa vlastníctva	TPE4	Autor remixu
PRIV	Rámec týkajúci sa súkromia	TPOS	Časť setu
PCNT	Počítadlo prehraní	TPUB	Vydavateľ
POPM	Ukazovateľ popularity	TRCK	Číslo skladby
POSS	Rámec týkajúci sa synchronizácie	TRDA	Dátumy nahrávky
RBUF	Odporúčaná veľkosť buffra	TRSN	Názov internetového rádia
RVAD	Doladenie hlasitosti	TRSO	Vlastník internetového rádia
RVRB	Spätná väzba	TSIZ	Veľkosť
SYLT	Synchronizované titulky	TSRC	ISRC
SYTC	Kódy pre synchronizáciu tempa	TSSE	Nastavenia použité na zakódovanie
TALB	Názov albumu/filmu/relácie	TYER	Rok
TBPM	BPM (údery za minútu)	TXXX	Textová informácia
TCOM	Skladateľ	UFID	Unikátny identifikátor súboru
TCON	Typ obsahu	USER	Podmienky použitia
TCOP	Správa o autorskom práve	USLT	Prepis titulkov
TDAT	Dátum	WCOM	Komerčná informácia
TDLY	Oneskorenie v zozname skladieb	WCOP	Informácia o autorskom práve
TENC	Meno osoby, ktorá obsah zakódovala	WOAF	Oficiálna stránka súboru
TEXT	Textár	WOAR	Oficiálna stránka interpreta
TFLT	Typ súboru	WOAS	Oficiálna stránka zdroju súboru
TIME	Čas	WORS	Oficiálna stránka internetového rádia
TIT1	Popis obsahu skupiny	WPAY	Informácia o platbe
TIT2	Názov skladby	WPUB	Oficiálna stránka vydavateľa
TIT3	Popis titulkov	WXXX	Užívateľom definovaná URL

3.1.2 Ostatné

- **Vorbis comment** – používaný v OGG Vorbis a FLAC formátoch
- **APE tag** – používaný v APE formáte, môže byť vložený do MP3 namiesto ID3

3.2 Obraz ako signál

Obraz je viacrozmerný signál. Je chápaný intuitívne ako obraz na sietnici ľudského oka alebo obraz nasnímaný kamerou. Môže byť modelovaný matematicky pomocou spojitkej skalárnej funkcie f dvoch alebo troch premenných, tomu sa hovorí obrazová funkcia. Obraz je teda popísaný obrazovou funkciou dvoch súradníc $f(x,y)$ v rovine. Obrazová funkcia troch premenných sa použije, buď keď sa plošné obrazy menia v čase t , tj. $f(x,y,t)$ alebo v prípade objemových obrazov $f(x,y,z)$. Hodnoty obrazovej funkcie odpovedajú niektorej meranej fyzikálnej veličine, napr. jas u obrazu z čiernobielej TV kamery, teploty u termovíznej kamery alebo schopnosti pohlcovať žiarenie v danom mieste objemu u röntgenového tomografu. Z obrazu môžeme zisťovať napr. veľkosť, polohu alebo priemerný jas vybraného objektu. [26]



Obrázok 4.1: Ukážka obrazovej funkcie [26]

3.2.1 Kompresia obrazov

Spochiava v redukcii množstva dát potrebných k reprezentácii obrazu. Spotrebované množstvo pamäte sa meria napr. v bajtoch. Kompresia sa používa pre uchovanie a prenos dát všeobecne.

Segmentácia objektov v obraze

Je potrebná interpretácia obrazu (metódy závislé na dátach). Dosahuje sa najvyšších kompresných pomerov. Nemožná spätná rekonštrukcia pôvodného obrazu.

Odstránenie redundantnej informácie

- Bezstratové metódy – umožňujú úplnú rekonštrukciu pôvodného signálu.
Formáty PNG, AVI (bez kompresie).
- Stratové metódy – umožňujú len čiastočnú rekonštrukciu pôvodného signálu.
Formáty DFT, FFT, JPEG, MPEG, MPS

3.3 Farebné modely

3.3.1 Grayscale

Grayscale obrázok je kvalitatívne slabší, ako farebný ekvivalent, no zaberá spravidla menší objem dát. Ak si pod jedným pixelom predstavíme jeden bajt, čo je 8 bitov, počet rôznych hodnôt tohto bajtu predstavuje 2^8 . Graficky sa tento pixel zobrazí ako farba zo stupnice šedej, rovnomerne rozdelenej na 256 farieb. Hodnota stupnice na jednom konci reprezentuje bielu farbu, na opačnom čiernu. Väčšinou má čierna farba hodnotu 0 a biela 255. Prevod do stupňov šedej nie je jednotný, pretože existuje viacero farebných modelov.

3.3.2 RGB

RGB je najznámejší farebný model, používa sa na zobrazovanie na obrazovke. Je rozdelený na 3 zložky: R – červená (red), G – zelená (green) a B – modrá (blue). Hodnota 0 predstavuje v každej zložke najtmavší odtieň, zložením všetkých troch núl dostaneme čiernu farbu. Opakom je biela farba, ktorá vznikne po zložení troch zložiek o hodnote 255. V pamäti zaberá jeden pixel 3 bajty, každá zložka reprezentuje 1 bajt.

Prevod z RGB do grayscale:

$$I = 0,30 \times R + 0,59 \times G + 0,11 \times B$$

kde I predstavuje hodnotu pixelu reprezentujúcu intenzitu na stupnici šedej.

3.3.3 YUV

YUV je farebný model používaný v televíznom vysielaní v normách PAL a HDTV. Jeden obrazový bod reprezentujú 3 zložky: Y – jasová zložka, U a V – farebné zložky. Jasová zložka má rozsah od 0 do 1 a farebné zložky majú rozsah od -0,5 do 0,5. [27]

Prevod z a do RGB:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,147 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1,137 \\ 1 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} * \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

3.3.4 YP_BP_R a YC_BC_R

YP_BP_R je farebný model používaný vo videokábloch s koncovkami typu cinch. YP_BP_R je analógová verzia modelu YC_BC_R, ktorý je určený pre digitálne video. Numericky sú ekvivalentné. Zložky: Y – jas, P_B/C_B – rozdiel medzi modrou farbou a jasom, P_R/C_R – rozdiel medzi červenou farbou a jasom.

Nie sú to absolútne farebné priestory, iba inak zakódovaný farebný model RGB. Od týchto 2 štandardov sa odvodzuje veľa iných rôznych kombinácií, používaných v rôznych TV systémoch, napríklad PAL, SECAM, NTSC, HDTV, H.262 a iné. Formát obrázku JPEG taktiež dovoľuje ich použitie. [28]

3.4 Detekcia hrán

Rôzne metódy transformácie obrazu sú väčšinou len matematické operácie, inak tomu nie je ani v mojej práci. Dve použité metódy sú rozobrané v tejto kapitole.

3.4.1 Konvolúcia

Konvolúcia je metóda, ktorá systematicky prechádza celý obraz a na výpočet novej hodnoty bodu využíva malé okolie O reprezentatívneho bodu. Táto hodnota je zapísaná do nového obrazu. Najčastejšie sa používajú obdĺžnikové masky s nepárnym počtom riadkov a stĺpcov, pretože v tom prípade môže reprezentatívny bod ležať v strede masky. Diskrétna konvolúcia má tvar:

$$g(x, y) = \sum \sum_{(m,n) \in O} h(x - m, y - n) f(m, n)$$

kde f je obrazová funkcia pôvodného obrazu, g je obrazová funkcia nového obrazu, h sa nazýva konvolučná maska alebo konvolučné jadro, h udáva koeficienty jednotlivých bodov v okolí O . Transformácie v lokálnom okolí bodu sa delia na dve skupiny:

Vyhľadzovanie – tieto metódy sa snažia potlačiť šum v obraze, ale rozostrejajú hrany.

- Ostrenie – detekcia hrán a čiar, ale zosilňuje šum.

Podľa matematických vlastností môžeme metódy predspracovania rozdeliť na

- Lineárne metódy – novú jasovú hodnotu bodu počítajú ako lineárnu kombináciu vstupných bodov, napríklad priemerovací filter
- Nelineárne metódy – berú do úvahy len body s určitými vlastnosťami, napríklad mediánový filter [29]

3.4.2 Sobelov operátor

Sobelov operátor je používaný v metódach spracovania obrazu, konkrétne pri hľadaní hrán. Operátor vyráta gradient v každom bode, čo je vlastne najväčší rozdiel bielej a čiernej farby v jeho okolí. Táto hodnota indikuje ako „drasticky“ sa zmenila svetlosť v okolí tohto bodu a podľa tejto intenzity môžeme usúdiť, či ide o okraj alebo nie. Je v ňom aplikovaný princíp konvolúcie.

Matematicky je to použitie dvoch matic o rozmere 3×3 , ktoré sú v konvolúcii s originálnym obrázkom. Pri priechode obrazovými bodmi treba vynechať okrajové body, pretože nemajú okraje a výpočet pre ne je neplatný. Výsledný obrázok je tak o 2 pixely kratší po výške aj po šírke. [30]

Výpočet pre každý bod

I	intenzita bodu v originálnom obrázku
M_x	prvá matica
M_y	druhá matica
G_x	gradient x-ovej osi
G_y	gradient y-ovej osi
G	magnitúda gradientu

Zadefinujeme matice.

$$M_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$$

$$M_y = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}$$

Vypočítame gradient x-ovej a y-ovej osi.

$$G_x = M_x * I$$

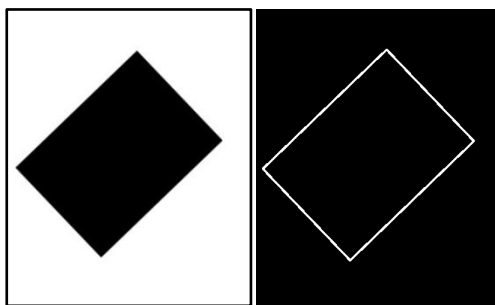
$$G_y = M_y * I$$

Vypočítame magnitúdu gradientu

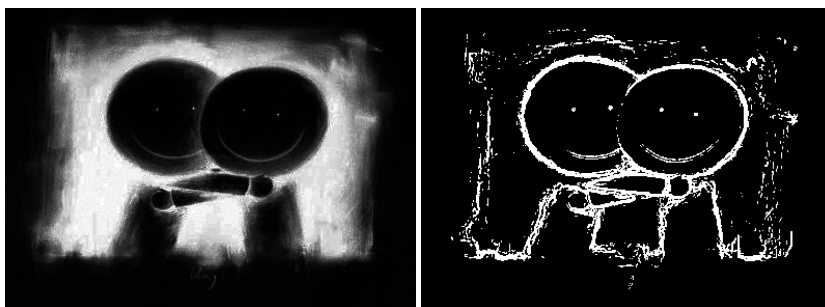
$$G = \sqrt{G_x^2 + G_y^2}$$

Hodnota magnitúdy gradientu nám vie pomôcť určiť, či je tento bod súčasťou okraja alebo nie. Túto hodnotu môžeme jednoducho prekresliť do obrázku, kde sa majú zvýrazniť okraje. Na takto vytvorenom obrázku sa dá pozorovať, do akej miery je oblasť kontrastná, pričom v najtmavších oblastiach sa intenzita bodov mení najmenej a naopak. Druhá možnosť je určiť si hraničnú hodnotu a porovnávaním gradientu s ňou určovať výslednú bielu (bod je súčasťou okraja) alebo čiernu (bod nie je súčasťou okraja) farbu.

Ukážky



Obrázok 4.2: Zľava: pôvodný obrázok, transformovaný Sobelovým operátorom



Obrázok 4.3: Komplikovanejší prípad

3.4.3 Prewittov operátor

Prewittov operátor je metóda detekcie hrán, ktorá počíta maximálnu odozvu zo série konvolúcií matíc, aby našla lokálnu orientáciu hrany pre každý pixel. Pre túto operáciu možno použiť rôzne matice. Séria 8 matíc je vyprodukovaná z jednej matice, ktorá rotuje a mení tak koeficienty. Hodnoty výstup obrazu orientácii ležia medzi 1 a 8, podľa toho, ktorá z 8 matíc dosiahla najvyššiu odozvu. Matematicky je to konvolúcia dvoch matíc a vstupného obrázku. [31]

Tieto matice vyzerajú nasledovne:

$$M_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix}$$

$$M_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix}$$

3.4.4 Cannyho hranový detektor

Cannyho zámer bol vynájdenie optimálneho detektoru hrán, v tomto prípade to znamená:

- dobrá detekcia – algoritmus by mal označiť čo najväčšie množstvo skutočných hrán
- dobrá lokalizácia – označené hrany majú byť čo najbližšie ku skutočnej hrane
- jednoznačná odozva – odozva na jednu hranu musí byť len jedna, nemôže dochádzať k hláseniu zdvojených hrán

Tento proces sa skladá zo 4 krokov:

1. Eliminácia šumu (napr. Gaussovým filtrom)

Dvojmerný variant Gaussovo normálneho rozdelenia je daná vzťahom:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}$$

kde x, y sú súradnice pixelu v obraze a σ je štandardná odchýlka rozdelenia (bežne =1 až 1,4). Výpočet je vhodné realizovať pomocou konvolúcie.

2. Určenie gradientu (prvá derivácia)

Využitie napríklad Sobelovho operátora.

3. Nájdenie lokálnych maxim (ztenšenie)

Cieľom tohto kroku je z hodnôt gradientu vybrať len lokálne maximá.

4. Eliminácia nevýznamných hrán (nastavenie prahu)

V tomto kroku si zvolíme minimálnu a maximálnu prahovú hodnotu. Ak sa hodnota gradientu daného pixelu bude pohybovať nad hornou hranicou, bude automaticky uznaný ako hranový. Ak sa táto hodnota bude pohybovať v rozmedzí prahových hodnôt, bude tento pixel uznaný ako hranový len v prípade, že leží vedľa pixelu, ktorý už bol takto uznaný. [32]

3.5 Houghova transformácia

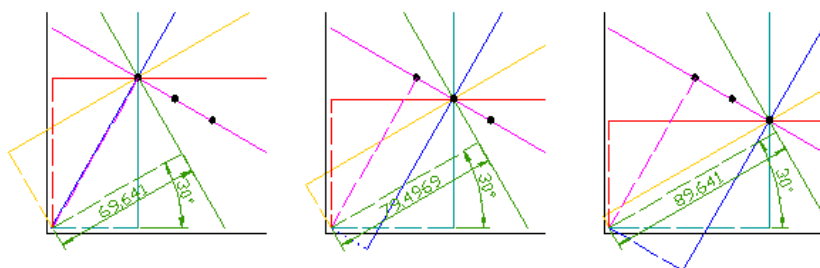
Houghova transformácia je technika použiteľná vtedy, keď treba detekovať objekty so známym tvarom hranice. Houghova transformácia môže detekovať rovné čiary aj krivky (hranice objektu), ak sú známe ich analytické vyjadrenia. Je robustná pri identifikácii zakrytých a zašumených objektov. V našom prípade ide o detekciu obdĺžnika alebo jemu podobného útvaru. Cieľom tohto algoritmu je vytvoriť tzv. *Houghov priestor*.

Je to matica bodov, kde vzdialenosť na jednej osi je ekvivalentom veľkosti uhla, ktorý zvierá bod, ktorý je uprostred pôvodného obrázku a kolmice na detekovanú hranu. Vzdialenosť na druhej osi je ekvivalentom vzdialenosti bodu na polpriamke od stredu obrázku, kde sa priamka, ktorej analytické vyjadrenie je totožné s hľadanou hranou, pretína s polpriamkou v uhle 90 stupňov.

Houghov priestor docielime tak, že nad obrázkom vytvoríme maticu čísel, ktorú možno nazvať *akumulátor*. Naplníme ich hodnotou 0, graficky znázornené to predstavuje čierny obraz. Pôvodný

obraz je ešte pred aplikáciou samotnej Houghovej transformácie vhodné transformovať pomocou niektorej z metódy detekcie hrán. V našom prípade sme použili Sobelov operátor. Uvažujme 3 body, kde bola detekovaná hrana („biely bod“) v obrázku, kde stred pôvodného obrázku je ľavý dolný roh.

1. Určíme si niekoľko uhlov, pod ktorými umiestnime priamky nad body (v našom prípade 0, 30, 60, 90 a 120 stupňov(Obr. 4.4).
2. Pre každú priamku vypočítame dĺžku úsečky, ktorá je na danú priamku kolmá a prechádza stredom pôvodného obrázku(Tab. 4.2).

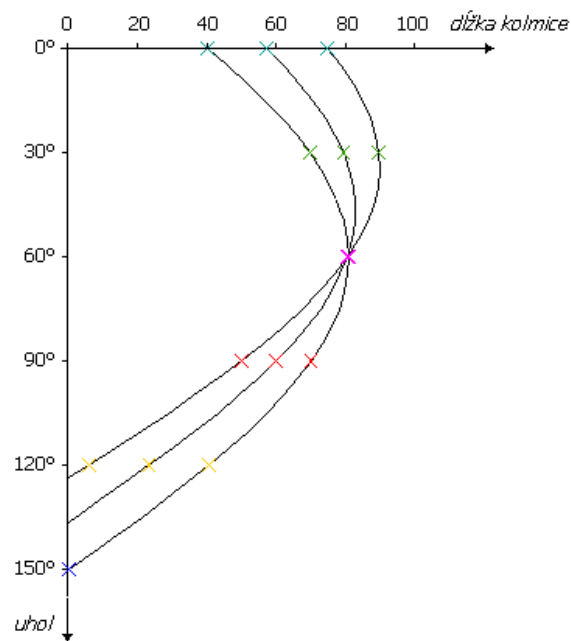


Obrázok 4.4: Ukážka bodov, priamok, vyrátaných úsečiek a ich dĺžok pri výpočte Houghovej transformácie, obrázok prevzatý z [33]

1. bod		2. bod		3. bod	
Uhol	Vzdialenosť	Uhol	Vzdialenosť	Uhol	Vzdialenosť
0°	40,0	0°	57,1	0°	74,6
30°	69,6	30°	79,5	30°	89,6
60°	81,2	60°	80,5	60°	80,6
90°	70,0	90°	60,0	90°	50,0
120°	40,6	120°	23,4	120°	6,0
150°	0,4	150°	-19,5	150°	-39,6

Tabuľka 4.2: Tabuľka vypočítaných hodnôt priechodu bodov pri výpočte Houghovej transformácie (zdrojové údaje pochádzajú z obrázku (Obr. 4.4))

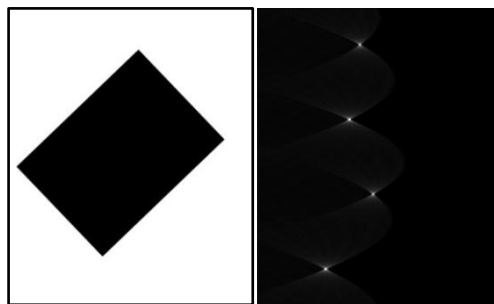
3. Pre každý bod vykreslíme do matice nad obrázkom krivku grafu, kde x-ová os predstavuje dĺžku kolmice na priamku reprezentujúcu hranu a y-ová os predstavuje uhol. Výsledok je matica, kde priesečník kriviek je bod, kde sa prekrývaním týchto kriviek „naakumulovala“ najväčšia hodnota. Tento bod sa v obrázku zobrazí ako najsvetlejší a reprezentuje detekovanú hranu(Obr. 4.5).



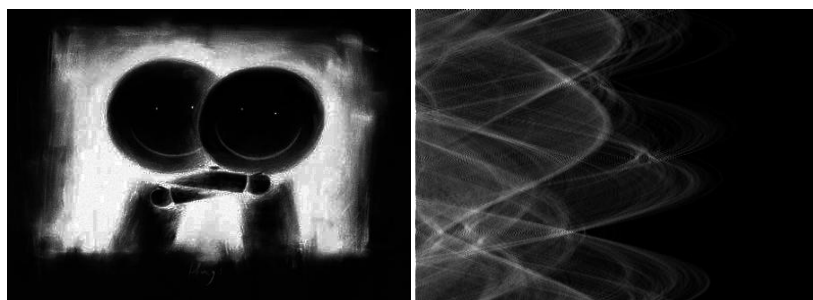
Obrázok 4.5: Nájdený bod (×), ktorý reprezentuje detekovanú hranu

4. Z obrázku možno usúdiť, že hrana sa nachádza na priamke, ktorá je kolmá na úsečku o dĺžke približne 80 obrazových bodov a ktorá zvierá so stredom pôvodného obrázku uhol 60 stupňov. Po algoritmickej stránke stačí nájsť najsvetlejší bod.

Ukážky



Obrázok 4.6: Zľava: pôvodný obrázok, transformovaný Houghov priestor



Obrázok 4.7: Komplikovanejší prípad

4 Požiadavky

Hlavná požiadavka bola navrhnutie užívateľského rozhrania pre editor ID3 tagov s ohľadom na automatizované hľadanie a korekciu tagov, pričom implementovaný editor má podporovať aspoň jeden stratový a jeden bezstratový hudobný formát. Aplikácia by mala umožniť používateľovi pohodlne spravovať svoju hudbu vrátane práce s obrazmi prebalov albumov.

5 Návrh aplikácie

Ako užívateľské rozhranie (pozn. ďalej GUI – Graphics User Interface) som zvolil komponent triedy `System.Windows.Forms.Form`. Je to klasické okno a zastávam názor, že atypické druhy okien sú pre väčšinu užívateľov skôr odpudivé. Toto okno je rozdelené na 2 logické časti. Prvú som nazval *Grabber*, obsluhuje sa v nej priebeh grabovacej činnosti. Automatizované hľadanie prebalu v poskytnutom obraze, či automatické hľadanie údajov sem taktiež patria. Druhá časť je nazvaná *Manager*, poskytuje užívateľovi možnosť upraviť údaje (tagy) v hudobných súboroch, prípadne zmeniť prebal a nakoniec spustiť prehrávanie. Pre zvýšenie abstrakcie problému som sa rozhodol použiť knižnice, ktoré sú zamerané na riešenie určitých častí implementácie. Všetky prvky, ktoré sú použité v aplikácii, sú použité podľa licencie, pod ktorou sú uverejnené.

5.1 Použité knižnice

5.1.1 LAME

Vývoj LAME začal v roku 1998. Pôvodne to bol patch na 8 Hz MP3 enkodéry, s projektom začal Mike Cheng. Neskôr bol kód prepísaný od začiatku, založený na zdrojoch enkodéra dist10. Za cieľ si položil zrýchlenie dist10 pri zachovaní kvality. Vývoj okolo vetvy, ktorej použite malo zostať pôvodné (patch), sa premenoval na Lame 2.0. V Lame 3.81 bol kód z dist10 zamenený meniac význam tejto vetvy tak, že už to ďalej nebol len patch. Projekt sa rýchlo stal tímovým projektom. Dnes je LAME považovaný ako najlepší MP3 enkodér pri stredných a vysokých prenosových rýchlostiach a vo VBR móde. Tento projekt je stále aktívne vyvíjaný. [34]

Funkcie LAME

- ✓ Veľa vylepšení v kvalite a rýchlosti oproti referenčnému softvéru podľa ISO
- ✓ Kódovanie do MPEG1, 2 a 2.5 (Layer III)
- ✓ CBR (konštantá prenosová rýchlosť), VBR a ABR (obe premenlivé prenos. rýchlosti)

- ✓ Kódovací engine môže byť skompilovaný v rôznych podobách
- ✓ Rýchlosť! Kóduje rýchlejšie než v reálnom čase na PentiumII 266 MHz v najvyššej kvalite
- ✓ A iné..[35]

Postup pri použití knižnice

1. Je potrebné vyplniť štruktúru *BE_CONFIG* a poslať ju parametrom do funkcie *beInitStream()*. Návrátová hodnota funkcie musí zodpovedať hodnote definovanej v konštante *BE_ERR_SUCCESSFUL*.
2. Následne treba rezervovať pamäť pre dáta. Minimálnu veľkosť sa dozvieme z *dwBufferSize*.
3. Voláme *beEncodeChunk()*, až kým sme nezakódovali všetko, čo je potrebné.
4. Voláme *beDeinitStream()*, to zaistí uvoľnenie pamäte ešte pred uzavretím streamu.
5. Zatvoríme stream zavolaním funkcie *beCloseStream()*.
6. Dodatočne, pri použití VBR kódovania, zavoláme funkciu *beWriteVBRHeader()*, ktorá vloží hlavičku INFO tagu, ktorý je rozšírením Xing VBR tagu. Informácia o použití tohto tagu musí byť taktiež zaznamenaná popredu v štruktúre *BE_CONFIG*. [36]

V prípade, že sa v adresári, v ktorom je uložená knižnica, nachádza súbor *lame_enc.ini*, enkodér doňho vpisuje debugovacie informácie.

V bakalárskej práci je táto knižnica využitá pri enkódovaní Audio CD do MP3 formátu.

5.1.2 TagLib#

TagLib# je .NET knižnica, ktorá umožňuje čítanie a zapisovanie metadát uložených v populárnych hudobných formátoch. Je založená na knižnici TagLib [37], ktorá je písaná v C++. Väčšina kódu je práve z nej a TagLib# je de facto port do C#, aby umožnil jednoduchšie, objektové použitie v prostredí .NET. [38] V súčasnej dobe podporuje ID3v1 a ID3v2 pre MP3 súbory, Ogg Vorbis comment, ID3 tagy and Vorbis commenty vo FLAC súboroch, MPC, Speex, WavPack and TrueAudio súbory.

V bakalárskej práci je táto knižnica využitá pri úpravách tagov pri grabovaní, aj pri menežovaní hudobných kolekcíí.

5.1.3 Interop (COM objekty)

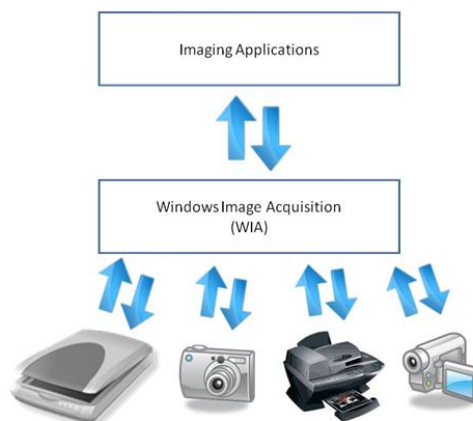
COM Interop poskytuje prístup k existujúcim COM komponentom bez nutnosti modifikácie pôvodného komponentu. [39]

5.1.3.1 Shell32

V bakalárskej práci je táto knižnica využitá pri získavaní prístupu k adresárom, konkrétne právam užívateľa pre daný adresár.

5.1.3.2 WIA (Windows Image Acquisition)

Windows Image Acquisition (WIA) je platforma pre získanie statického obrazu, funkčná v operačných systémoch Microsoft Windows, počínajúc Windows Me a Windows XP. [40]



Obrázok 6.1: Demonštrácia WIA frameworku. [40]

V bakalárskej práci je táto knižnica využitá pri získavaní obrazu z externého zariadenia pri získaní obrazu prebalu CD z web kamery, skenera, digitálneho fotoaparátu, či iného kompatibilného zdroja.

5.1.4 Ripper, Yeti.mmedia a Yeti.mp3

Sada knižníc poskytujúcich API na objektovú prácu s LAME enkodérom a CD-ROM mechanikami a ich ekvivalentmi. [41] V bakalárskej sú tieto knižnice použité na zvýšenie abstrakcie návrhu. Poskytujú objektovo orientovaný prístup k nemenežovanej knižnici lame_enc.dll a k API funkciám systému Windows.

5.2 Grafické prvky

Všetky ikony, ktoré sa nachádzajú v aplikácii, pochádzajú z portálu Iconfinder.com. Tento projekt jednoduchým a efektívnym spôsobom poskytuje vysoko kvalitné ikony pre webdizajnérov a vývojárov. [42]

5.3 Vývojový nástroj

K vývoju aplikácie som použil Microsoft Visual Studio 2008 (verzia 9.0.21022.8). Visual Studio je vhodné pre vývojárov, dizajnérov, architektov, testerov, projektových manažérov a ďalších účastníkov životného cyklu vývoja softvérových aplikácií. [43] Ako vývojový jazyk som si zvolil C#. Je objektovo orientovaný a moderný. Počas „surfovania“ po internete, keď som hľadal čo najviac informácií, týkajúcich sa môjho problému a riešenia, som usúdil, že je najrozšírenejší spomedzi rodiny .NET jazykov. Veľmi podobný jazyk je aj .NET C++, avšak nemá veľkú podporu v komunite.

5.3.1 Technológia .NET

.NET („dotnet“ podľa anglického dot NET = bodka NET, NET pochádza z network, sieť) je zastrešujúci názov pre súbor technológií v softvérových produktoch, ktoré tvoria celú platformu, ktorá je dostupná nielen pre Web, Windows a Pocket PC. Common Language Infrastructure je štandardizovaná špecifikácia jadra .NET. Základným komponentom je Microsoft .NET Framework, prostredie potrebné pre beh aplikácií a ponúkajúce ako spúšťačie rozhranie, tak potrebné knižnice. Pre vývoj .NET aplikácií vydal Microsoft Visual Studio .NET. [44]

5.4 Implementácia

Keďže moja práca spája hudobné a obrazové metódy, rozhodol som sa rozdeliť zdrojový kód do 2 hlavných tried, `mainForm.cs` a `imageRecognition.cs`. Dielčie funkcie a skupiny funkcií sú kategoricky rozdelené do regiónov v rámci každej triedy.

5.4.1 `mainForm.cs`

Táto trieda obsluhuje funkcie potrebné na prácu s hudobnou časťou, pričom je rozdelená do 6 regiónov. V grafickej podobe sa táto trieda delí na 2 módy – *Grabber* a *Manager*, prepínajú sa v hornej časti GUI pomocou tlačidla (Obr. 6.2).



Obrázok 6.2: Prepínač módu.

5.4.1.1 Constants a Globals

Tieto regióny sú umiestnené v úvodnej časti zdrojového kódu, obsahujú použité konštanty a premenné, ktoré sú hlavne stavového charakteru. Konštanty predstavujú určité popisy v angličtine. Pôvodný (rozšírený) plán bol implementovať použitie všetkých reťazcov pomocou konštánt a premenných typu `string`, a tak umožniť ľahké implementovanie viacerých jazykov. Keďže však tento zámer nie je predmetom práce, pôvodné konštanty ostali nezmenené, no počas písania som sa rozhodol zamerať sa na pre program nevyhnutnejšie algoritmy.

5.4.1.2 Init

Obsahuje konštruktor, v ktorom sa inicializujú komponenty pomocou vygenerovaného, aj mnou písaného kódu.

5.4.1.3 Common

Obsahuje funkcie potrebné pre konverziu dát a rôzne pomocné funkcie, ktoré sú alebo môžu byť použité v celej aplikácii. Takisto sú tu funkcie obsluhujúce prepínanie medzi zvolenými módmi.

Pre príklad uvediem napríklad funkcie:

`String validFName(String str)` – funkcia, ktorej návratová hodnota je reťazec v podobe opraveného názvu súboru, tak, aby bol platný.

`void setMode(Byte mode)` – funkcia, ktorá aktualizuje GUI, tak, že je zobrazený buď mód *Grabber* alebo *Manager*, podľa toho, ktorý mód vyberie používateľ v hornej časti GUI.

5.4.1.4 Grabber

Všetky funkcie dôležité pre kompletný grabovací proces sú umiestnené v tomto regióne. Tento proces som sa najprv rozhodol vytvoriť v štýle dialógu typu „wizard“. Avšak takýto druh dialógu umožňuje prepínanie len na nasledujúce alebo predchádzajúce okno. Rozhodol som sa teda použiť komponent triedy `TabControl`. Ten umožňuje jednoduché a ľubovoľné prepínanie medzi oknami (pozn. ďalej tabmi).

Celý proces som rozdelil na 5 častí, ktoré sú zoradené do istej mieri chronologicky, podľa toho, v akom poradí by úkon grabovania mal byť uskutočnený. Zoznam týchto častí:

- Select drive
- Set artwork
- Fill album data
- Choose format
- Finalize

Select drive

V tomto tabe má užívateľ možnosť vybrať si zdrojovú mechaniku pomocou `ComboBoxu`.

V nej vidno písmeno jednotky, model a bus. V dolnej časti je zobrazený TOC (informácie o médiu – Table of Contents). Obsahuje zoznam skladieb s veľkosťami, ktoré na disku zaberajú. Ak sa na disku nachádza aspoň jedna audio stopa, je toto CD načítané do inštancie objektu typu `CDDrive` z knižnice `Ripper`. Následne sú oživené ďalšie taby tam, kde to má význam. Napríklad dopĺňanie informácií o albume nemá význam, pokiaľ nemáme zdrojové dáta. Naopak výber obrázku prebalu môže prebiehať nezávisle.



Obrázok 6.3: Zobrazené TOC.

Set artwork

Tu používateľ nastaví použitý obraz prebalu. Má na výber import zo súboru alebo z externého zariadenia. Pri použití externého zdroja, musí byť prebal z obrazu extrahovaný, a teda musí nasledovať dialóg s implementovanými metódami transformácie obrazu (obsluhuje trieda `imageRecognition.cs`). Pri importovaní zo súboru je použitie tohto dialógu voliteľné (pomocou

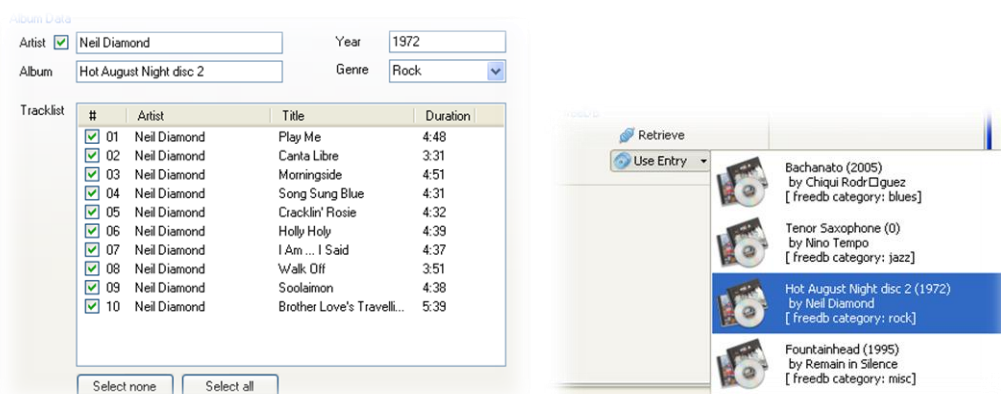
CheckBoxu). Obráz je následne v tomto tabe zobrazený s možnosťou zmazania. Importovanie zo z externého zdroja zabezpečuje zaimplementovaná knižnica WIA.

Fill album data

V tomto tabe má používateľ možnosť upraviť dáta albumu. V hornej časti som umiestnil textové polia, v ktorých možno upraviť názov interpreta, albumu, žáner a rok albumu. Názov interpreta sa dá označiť ako predvolený pre všetky skladby, pokiaľ disk nie je kompilácia. Na pole so žánrom som využil ComboBox s možnosťou editovania, ktorý je naplnený zoznamom žánrov, ktorý nájdeme v špecifikáciách ID3 tagu. Zoznam skladieb som umiestnil do komponenty ListBox. Pre každú skladbu je zobrazená dĺžka trvania a možno nastaviť meno interpreta a názov skladby.

Automatizácia je v tomto tabe zastúpená v podobe automatického načítania dát o albume zo serveru služby freedb. Na server pošleme požiadavku, ktorá obsahuje *Disc-ID*. Je to 8-bajtové číslo, ktoré je vygenerované z údajov v TOC podľa algoritmu na stránkach freedb[27]. Toto ID nám poskytne metóda `GetCDDBDiskID()` z inštalácie objektu triedy `CDDrive`, ktorú sme aktualizovali v prvom tabe pri výbere zdrojovej mechaniky. Použitím tlačidla *Retrieve* získame dáta pomocou funkcie `string retrieveFreeDBtext(String category, String discid)` a naplníme tlačidlo *UseEntry*, ktoré je typu `ToolStripDropDownButton` základnými informáciami o albume tak, aby pre používateľa bolo jednoduché identifikovať vybraný album. Po kliknutí na tlačítko *UseEntry* sa následne vyroluje menu, z ktorého kliknutím používateľ album vyberie a všetky údaje sa do tabu takto vyplnia.

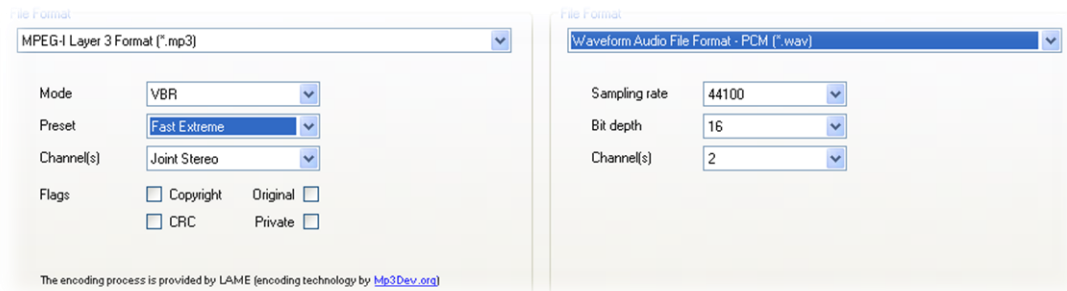
Zmena informácií o skladbách sa uskutočňuje v pravom hornom rohu, kde môže používateľ intuitívne používať šípky na klávesnici (▼▲) pre prechádzanie po skladbách a okamžitú úpravu údajov. Tieto sú uložené do zoznamu skladieb tlačением klávesy *ENTER* (prípadne kliknutím na tlačidlo *OK*).



Obrázok 6.4: Zobrazené údaje o albume a použitie freedb.

Choose file format

Pomocou ComboBoxu si tu používateľ vyberie formát súboru, do ktorého chce hudbu uložiť. Na výber je formát MP3 alebo WAV. Pre každý formát sú dostupné vlastné nastavenia, ktoré sú umiestnené na komponente Panel. Tento má v oboch prípadoch nastavený okraj na žiadny (vlastnosť `Border`) a dokovanie na vyplnenie (vlastnosť `Dock`). Menením indexu v ComboBoxe sa len mení viditeľnosť oboch panelov naopak.



Obrázok 6.5: Nastavenia formátov.

Ďalej je na tomto tabe možnosť nastavenia formátu názvu súboru v podobe TextBoxu. V ňom reprezentujú reťazce zložené z dvoch znakov jednotlivé položky, ktoré sú vyplnené v časti *Fill album data*. Konkrétne reťazce a položky, ktoré reprezentujú:

- %1 – číslo skladby
- %2 – meno umelca
- %3 – názov skladby
- %4 – názov albumu

Finalize

V tomto tabe sú zobrazené tlačidlo na spustenie a zrušenie grabovacieho procesu. Takisto tu má užívateľ možnosť sledovať priebeh grabovania prostredníctvom komponenty `ProgressBar`, pre práve grabovanú skladbu, aj pre celkový priebeh. Vybráním príslušného adresára a stlačením tlačidla *Grab It!* Sa spustí grabovací proces. Ten pozostáva z viacerých fáz a obsluhuje ich funkcia `grabCD`:



1. Príprava GUI – zresetovanie `ProgressBar`ov, umrtvenie tabov a pod.
2. Príprava formátu názvu súboru – zámena podreťazcov „%1“ -> „{0}“. Na vstupný reťazec som viacnásobne použil metódu `String.Replace` a tým som dostal reťazec použiteľný ako 1. parameter v metóde `String.Format`, kde tento reťazec predstavuje formát s miestami, kam sa vložia reťazce položiek albumu. Následne bude použité pri jednotlivých súboroch.
3. Ak používateľ vybral prebal albumu, tento sa uloží do súboru.
4. Uzamkneme mechaniku pomocou metódy `CDDrive.LockCD()`
5. Prechádzame skladby, ktoré používateľ označil.
6. Pomocou metódy `String.Format` a funkcie `validFName` získame názov súboru.

7. Enkódujeme do WAV formátu, pričom využívame triedy `WaveFormat`, ktorú inicializujeme s nastaveniami, ktoré používateľ vybral v GUI. Vytvoríme zapisovací `Stream` spustíme čítanie dát z CD pomocou metódy `CDRdrive.ReadTrack`.
8. V prípade grabovania do MP3 aktualizujeme štruktúru `BE_CONFIG` pomocou funkcie `updateBECONFIG()`. Vytvoríme `Stream` na zapisovanie do MP3 a vytvoríme mp3 súbor.
9. Zapišeme metadáta do súboru, teda tagy.
10. Odstránime dáta, ktoré sú nepotrebné.
11. Uvoľníme `Stream(y)`.
12. Odomkneme mechaniku pomocou metódy `CDDrive.LUnockCD()`.
13. Aktualizujeme GUI.

5.4.1.5 Manager

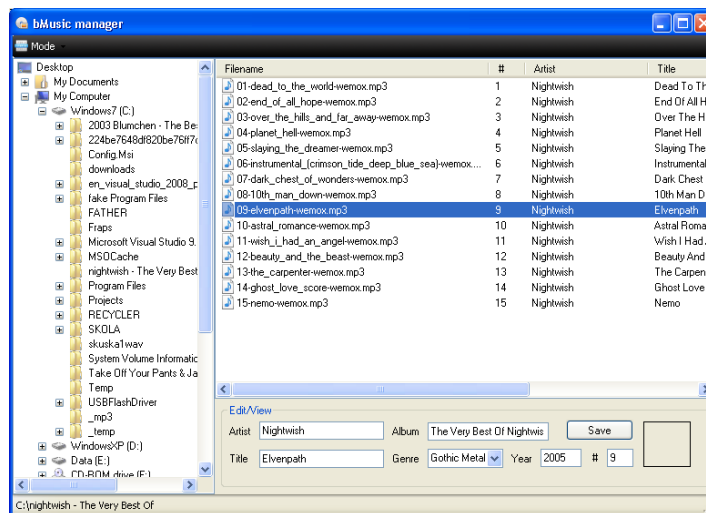
Tento región obsahuje funkcie, ktoré sú využité v móde *Manager*. Sú rozdelené do podregiónov *DirTree* (ľavá časť GUI) a *FileView* (pravá časť GUI).

DirTree

Na výrobu štandardného „stromu“ štruktúry adresárov som použil komponent `TreeView`. Tento strom sa snaží napodobniť strom, ktorý je použitý v operačných systémoch Microsoft Windows. To znamená, že prvý je zobrazený adresár „Moje dokumenty“, potom „Tento počítač“, ktorý je predvolene rozkliknutý a zobrazené sú disky a mechaniky, nasledujú adresáre z „Plochy“. Objekty v strome sú inicializované pomocou funkcie `updateTree()`. Pred rozkliknutím (event `BeforeExpand`) je vyplnená vždy hlbšia úroveň adresárov. Ikona  v „Prieskumníkovi“ signalizuje prítomnosť podadresárov. Podpora tejto funkcie vyžaduje vždy „prejsť“ aj podadresáre, či sú alebo nie sú prázdne. Ak obsahujú podadresár, vloží sa do takéhoto podadresára tzv. „dummy node“. Je to neexistujúci subadresár, ktorý sa však nikdy nezobrazí, ale práve umožní zobrazenie ikony . Po označení validného adresára sa cesta prepíše do spodnej časti GUI, do `StatusBaru`. Zmena tejto hodnoty vyvolá aktualizáciu pravej časti GUI.

FileView

Tento región obsahuje funkcie pre úpravu tagov a zobrazenia podporovaných formátov v zozname súborov v pravej časti GUI. Po aktualizovaní cesty adresára sa v zozname súborov vypíšu podporované súbory aj s metadátami. Po označení súboru sa v dolnej časti dajú upravovať dáta tohto, prípadne viacerých súborov. Dvojklik spustí prehrávanie v prehrávači, ktorý má používateľ nastavený ako predvolený. Úprava dát sa potvrdí tlačidlom *Save*.

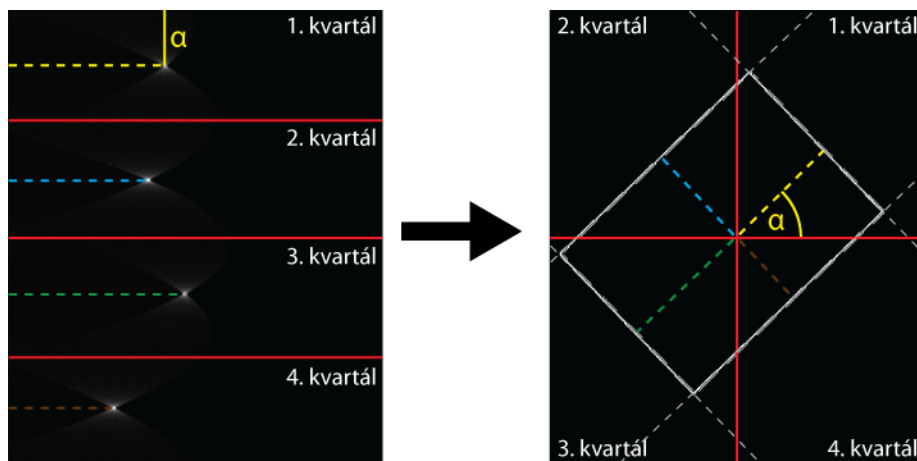


Obrázok 6.6: Manager mód.

5.4.2 imageRecognition.cs

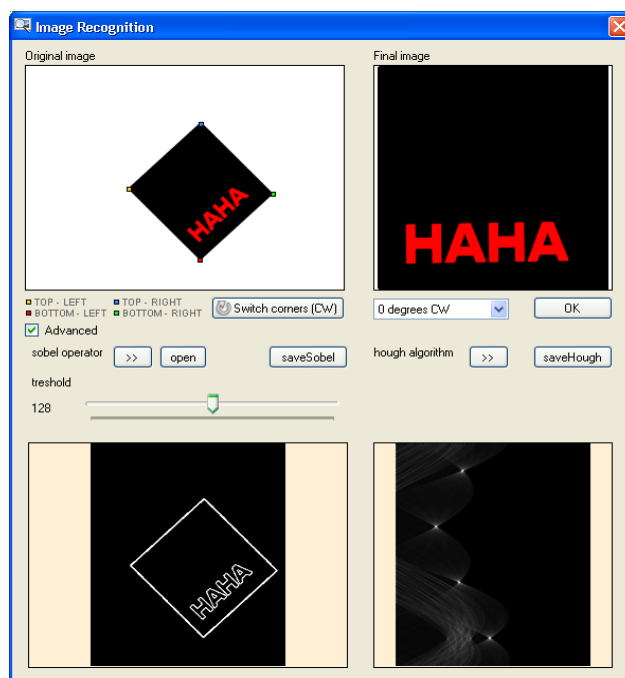
Táto trieda neobsahuje regióny. Z hlavného okna je do inštancie objektu tejto triedy uložený zdrojový obrázok. Po udalosti Shown inštancie hlavného komponentu Form, najprv obraz zmenšíme, aby sme šetrili výpočtový čas procesora, pričom pôvodnú bitmapu s pomerom zväčšenia uchováme v podobe globálnych premenných. Nasleduje spustenie funkcie `doSobel()`, ktorá vygeneruje nový obraz, v ktorom sú zobrazené nájdené hrany. Pokračuje funkcia `doHough()`, ktorá transformuje tento obraz na ďalší obraz, ktorý reprezentuje Houghov priestor.

Vychádzal som z faktu, že hľadaný prebal prechádza stredom obrázku. Takto sa v každom kvadrante nachádza jedna strana obalu. V Houghovom priestore to znamená, že celý obraz možno rozdeliť na 4 rovnako vysoké obdĺžniky, pričom každý reprezentuje jeden kvadrant. Keďže hľadáme teoreticky 1 stranu prebalu v 1 kvadrante, z toho vyplýva, že v každom z týchto obdĺžnikov stačí nájsť najsvetlejší bod. Z týchto bodov reprezentujúcich priamky následne vyrátam priesečníky s okrajmi obrázky ($x = 0$, $y = 0$, $x = Width(orig)$, $y = Height(orig)$). Vždy z dvoch priamok použijem tieto 2 body a pomocou funkcie `intersectionPoint(Point startPointA, ..., Point endPointB)` vyrátam priesečníky, čo sú vlastne rohy prebalu CD. Jej vstupné parametre sú 2 body typu `Point`, začiatkový a konečný, pričom vracia priesečník opäť typu `Point`. Podľa pomeru zmenšení tieto body umiestnim do GUI v podobe 4 malých bodov. Celý tento proces je indikovaný `ProgressBar`om.



Obrázok 6.7: Rekonštrukcia priamok reprezentujúce strany prebalu z Houghovho priestoru

Funkcia `extractImage()` podľa umiestnení týchto bodov spätne vyrába skutočné súradnice a extrahuje finálny obraz. Pomocou udalostí rodiny `Mouse` sa nimi dá manipulovať, pričom je na nich zobrazený kurzor v tvare kríža pre presnejšiu orientáciu. Extrakcia prebieha transformáciami nad objektom typu `Graphics`. Najskôr je obraz natočený tak, aby spodná hrana bola rovnobežne s osou x . Sú tu využité princípy z trigonometrie, následne treba bitmapu posunúť ku rohu a orezať. Po stlačení tlačidla *OK* je tento vložený do hlavného okna. Prípadne môžeme zobrazit' rozšírený dialóg po zaškrtnutí `CheckBoxu` s menovkou *Advanced*. Tam je umožnené vkladanie vlastnej definície bodov (obrázok v podobe Houghovho priestoru), prípadne je možné meniť prah citlivosti u Sobelovho operátora.

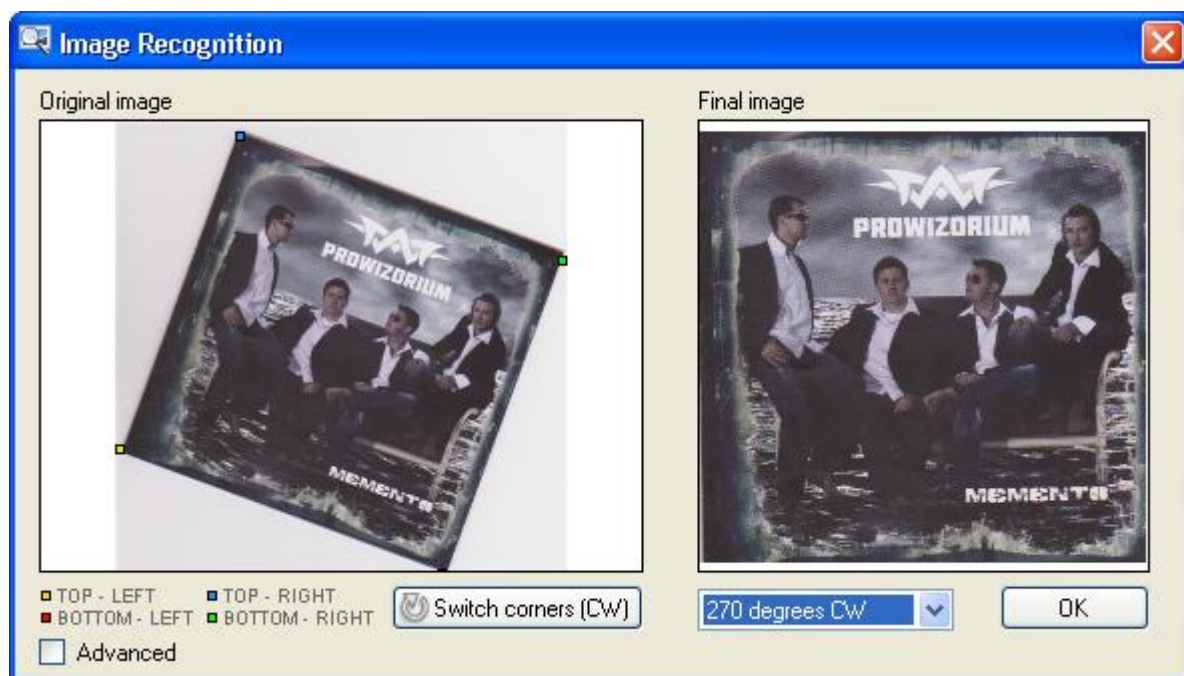


Obrázok 6.8: Rozšírený dialóg na detekovanie prebalu.

5.5 Test aplikácie

Aplikáciu som otestoval na grabovanie kompresného aj nekompresného formátu. Tieto fungujú a v prehrávačoch sa správajú štandardne. Problém nastáva pri poškodených CD, pričom počas čítania takéhoto disku je ošetrovaná výnimka a disk sa neskopíruje. Rozpoznávanie obrazu nie je 100%-tné. Závisí od vstupného obrázku, ako si algoritmus s týmto poradí. Ak obsahuje obal CD veľké tieň, prípadne ak sa na obraze nachádza prílišné množstvo hrán, môže nastať, že algoritmus deteguje inú hranu. Kvôli tomuto nedostatku som implementoval dodatočné manuálne nastavenie rohov prebalu. Každý roh funguje na princípe drag'n'drop, takže používateľ si tieto rohy pohodlne manuálne doladí a pri pustení tlačidla myši hneď vidí výsledok.

Skúška detekcie pri skenovaní pri 150 dpi



Obrázok 6.9: Ukážka detekovania v reálnom prostredí.

Manuálne som musel nastaviť otočenie o 270° v smere hodinových ručičiek. Body sú zobrazené tak, ako ich detekoval program. Občas sa stávalo, že jeden, či viacej bodov uskočilo.

6 Záver

Predmetom práce bola demonštrácia grafického prostredia, kde by sa dali upravovať tagy hudobných súborov a zároveň by boli využité metódy spracovania obrazu.

Obraz v hudbe predstavujú prebaly CD, vhodné spojenie predstavuje implementácia grabovania Audio CD do hudobných súborov pre neskoršiu editáciu a prehrávanie. Aplikácia je rozdelená na 2 časti „manager“ a „grabber“, a tým sa logicky oddeľuje grabovanie a editovanie. V grabovacej časti má užívateľ možnosť využiť rozpoznávanie obrazu. Pri ukladaní súborov sa totiž ukladá aj obraz – prebal disku.

Možné rozšírenia

- Integrovaný prehrávač
- Podpora viacerých formátov pri grabovaní.
Formáty ako OGG, FLAC, AAC ,WMA alebo APE by určite prispeli k prípadnej rozšírenosti.
- Podpora OCR a kvalitnejšie rozpoznávacie metódy
Album by takto mohol byť identifikovaný už podľa prebalu, čo by zvýšilo mieru automatizácie.
- Automatické sťahovanie prebalov z webových zdrojov
- Rôzne zdroje hudby
 - FTP
 - Web (s vyhľadávaním)
 - Torrent (s vyhľadávaním)
 - Line-In (nahrávanie pomocou audio vstupu)

Použitie

Táto aplikácia má slúžiť ako náhrada iných grabovacích softvérových riešení, pričom sa zameriava na určitú mieru automatizácie a jednoduchosti spojenú s možnosťou dodatočnej úpravy tagov. Mieru automatizácie reprezentuje automatické dopĺňanie údajov albumu zo služby freedb a rozpoznávanie prebalu v obraze.

Literatúra

- [1] Hudba na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 29. február 2004, 13. máj 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://sk.wikipedia.org/wiki/Hudba>>.
- [2] DANIEL, Eric D.; MEE, C. Dennis; CLARK, Mark H. *Magnetic Recording: The First 100 Years*. Chicago : Wiley-IEEE Press, 1999. 370 s. ISBN 0780347099.
- [3] *WhatIs.com* [online]. 2000-20-05 [cit. 2010-05-15]. What is Red Book?. Dostupné z WWW: <http://searchstorage.techtarget.com/sDefinition/0,,sid5_gci503642,00.html>.
- [4] EMERAN, Riyad. Trusted Reviews [online]. 2008-04-08 [cit. 2010-05-15]. Music On The Move. Dostupné z WWW: <<http://www.trustedreviews.com/mp3/review/2008/04/08/Music-On-The-Move/p4>>.
- [5] DVD na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 8. september 2004, 13. máj 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://sk.wikipedia.org/wiki/DVD>>.
- [6] *Blu-ray.com* [online]. 2010 [cit. 2010-05-15]. Blu-ray FAQ. Dostupné z WWW: <<http://www.blu-ray.com/faq/>>.
- [7] Red Book (audio Compact Disc standard) na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 26 February 2002, 10 April 2010 [cit. 2010-05-15]. Dostupné z WWW: <http://en.wikipedia.org/wiki/Red_Book_%28audio_Compact_Disc_standard%29>.
- [8] *Apple* [online]. CUPERTINO, California : 2010-02-25 [cit. 2010-05-15]. iTunes Store Tops 10 Billion Songs Sold. Dostupné z WWW: <<http://www.apple.com/pr/library/2010/02/25itunes.html>>.
- [9] *BBC NEWS* [online]. 2009-01-07 [cit. 2010-05-15]. Apple to end music restrictions. Dostupné z WWW: <<http://news.bbc.co.uk/2/hi/technology/7813527.stm>>.
- [10] *Apple* [online]. - [cit. 2010-05-15]. Apple iTunes. Dostupné z WWW: <<http://www.apple.com/itunes/>>.
- [11] HULÁN, Radek. *MyEgo.cz* [online]. 2006 [cit. 2010-05-15]. Stahovat si hudbu pro vlastní potřebu je legální. Dostupné z WWW: <<http://myego.cz/item/stahovat-si-hudbu-pro-vlastni-potrebu-je-legalni>>.
- [12] Peer-to-peer na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 1. apríl 2005, 6. máj 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://sk.wikipedia.org/wiki/Peer-to-peer>>.
- [13] WAV na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 26. 10. 2006, 5. 4. 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://cs.wikipedia.org/wiki/WAV>>.
- [14] *Digital Preservation (Library of Congress)* [online]. - [cit. 2010-05-15]. WAVE Audio File Format. Dostupné z WWW: <<http://www.digitalpreservation.gov/formats/fdd/fdd000001.shtml>>.
- [15] *FLAC - Free Lossless Audio Codec* [online]. 2008 [cit. 2010-05-15]. FLAC - comparison. Dostupné z WWW: <<http://flac.sourceforge.net/comparison.html>>.
- [16] Monkey's Audio na *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 21. 10. 2007, 13. 4. 2010 [cit. 2010-05-15]. Dostupné z WWW: <http://cs.wikipedia.org/wiki/Monkey%27s_Audio>.
- [17] MP3 In *Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 25. máj 2006, 7. máj 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://sk.wikipedia.org/wiki/MP3>>.

- [18] *RFC 3533 - The Ogg Encapsulation Format Version 0* [online]. 2003 [cit. 2010-05-15]. IETF Tools. Dostupné z WWW: <<http://tools.ietf.org/html/rfc3533>>.
- [19] *AAC na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 23. 8. 2006, 13. 4. 2010 [cit. 2010-05-15]. Dostupné z WWW: <<http://cs.wikipedia.org/wiki/AAC>>.
- [20] *Windows Media Audio na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 26. 8. 2006, 13. 4. 2010 [cit. 2010-05-15]. Dostupné z WWW: <http://cs.wikipedia.org/wiki/Windows_Media_Audio>.
- [21] *Freedb.org* [online]. - [cit. 2010-05-15]. Freedb.org > FAQ. Dostupné z WWW: <<http://www.freedb.org/en/faq.3.html#11>>.
- [22] *Gracenote* [online]. - [cit. 2010-05-15]. Gracenote: MusicID. Dostupné z WWW: <http://www.gracenote.com/business_solutions/music_id/>.
- [23] *Cdcovers.cc / World's Largest CD Covers and DVD Covers Album Art Archive* [online]. 2010 [cit. 2010-05-15]. Cdcovers.cc / World's Largest CD Covers and DVD Covers Album Art Archive. Dostupné z WWW: <<http://www.cdcovers.cc/>>.
- [24] *ID3.org* [online]. 2006-10-29 [cit. 2010-05-15]. ID3v1 - ID3.org. Dostupné z WWW: <<http://www.id3.org/ID3v1>>.
- [25] *ID3.org* [online]. 2006-10-29 [cit. 2010-05-15]. id3v2.3.0 - ID3.org. Dostupné z WWW: <<http://www.id3.org/id3v2.3.0>>.
- [26] *Webzdarma.cz* [online]. 2006 [cit. 2010-05-17]. 2006_K44.pdf. Dostupné z WWW: <http://www.bakal06.chytrak.cz/2006_K44.pdf>.
- [27] *YUV na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 29 September 2002, 6 May 2010 [cit. 2010-05-17]. Dostupné z WWW: <<http://en.wikipedia.org/wiki/YUV>>.
- [28] *YPbPr na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 15 July 2004, 8 May 2010 [cit. 2010-05-17]. Dostupné z WWW: <<http://en.wikipedia.org/wiki/YPbPr>>.
- [29] *Interaktívna učebnica spracovania obrazu - Katedra aplikovanej informatiky, FMFI UK Bratislava* [online]. - [cit. 2010-05-17]. Predspracovanie. Dostupné z WWW: <<http://dip.sccg.sk/predspra/predspra.htm>>.
- [30] VEELAERT, Peter; TEELLEN, Kristof. Adaptive and optimal difference operators in image processing [online]. Ghent University/IBBT : Elsevier, 2008 [cit. 2010-05-15]. Dostupné z WWW: <linkinghub.elsevier.com/retrieve/pii/S0031320308004810>.
- [31] *Prewitt na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 4 June 2005, 26 March 2010 [cit. 2010-05-17]. Dostupné z WWW: <<http://en.wikipedia.org/wiki/Prewitt>>.
- [32] *Cannyho hranový detektor na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 21. 6. 2006, 6. 10. 2009 [cit. 2010-05-17]. Dostupné z WWW: <http://cs.wikipedia.org/wiki/Cannyho_hranov%C3%BD_detektor>.
- [33] *Hough transform na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 18 January 2004, 10 May 2010 [cit. 2010-05-15]. Dostupné z WWW: <http://en.wikipedia.org/wiki/Hough_transform>.
- [34] *LAME MP3 Encoder* [online]. March 2010 [cit. 2010-05-15]. LAME MP3 Encoder. Dostupné z WWW: <<http://lame.sourceforge.net/>>.
- [35] *LAME MP3 Encoder* [online]. - [cit. 2010-05-15]. LAME MP3 Encoder :: About. Dostupné z WWW: <<http://lame.sourceforge.net/about.php>>

- [36] FABER, Albert L. LAMEOnJ - *The LAME symmetric Java MP3 encoder API* [online]. [cit. 2010-05-15]. Lame-. Dostupné z WWW: <<http://openinnowhere.sourceforge.net/lameonj/LameDLLInterface.htm>>.
- [37] *TagLib* [online]. - [cit. 2010-05-15]. TagLib. Dostupné z WWW: <<http://developer.kde.org/~wheeler/taglib.html>>.
- [38] *Developer Community* [online]. 2009 [cit. 2010-05-15]. TagLib Sharp - Developer Community. Dostupné z WWW: <http://developer.novell.com/wiki/index.php/TagLib_Sharp>.
- [39] *MSDN: Microsoft Development, MSDN Subscriptions, Resources, and More* [online]. [cit. 2010-05-15]. COM Interop Part 1: C# Client Tutorial (C#). Dostupné z WWW: <<http://msdn.microsoft.com/en-us/library/aa645736%28VS.71%29.aspx>>.
- [40] *MSDN: Microsoft Development, MSDN Subscriptions, Resources, and More* [online]. [cit. 2010-05-16]. Windows Image Acquisition (WIA). Dostupné z WWW: <<http://msdn.microsoft.com/en-us/library/ms630368%28VS.85%29.aspx>>.
- [41] *Your Development Resource - CodeProject* [online]. - [cit. 2010-05-16]. Articles by Idael Cardoso (Articles: 5). Dostupné z WWW: <<http://www.codeproject.com/script/Articles/MemberArticles.aspx?amid=325803>>.
- [42] *Iconfinder | Download free icons* [online]. - [cit. 2010-05-16]. Iconfinder | Download free icons. Dostupné z WWW: <<http://www.iconfinder.com/>>.
- [43] *MSDN: Microsoft Development, MSDN Subscriptions, Resources, and More* [online]. 2008 [cit. 2010-05-16]. Microsoft Visual Studio. Dostupné z WWW: <<http://www.microsoft.com/slovakia/msdn/produkty/vstudio/default.msp>>.
- [44] *.NET na Wikipedia : the free encyclopedia* [online]. St. Petersburg (Florida) : Wikipedia Foundation, 14. december 2008 , 28. apríl 2010 [cit. 2010-05-16]. Dostupné z WWW: <<http://sk.wikipedia.org/wiki/.NET>>.

Zoznam príloh

Príloha 1. Informácie k spusteniu

Príloha 2. CD

Príloha 1. Informácie k spusteniu

- Keďže je tento projekt implementovaný v jazyku C#, nevyhnutné mať v systéme nainštalovaný .NET Framework Redistributable Package. V tomto prípade je potrebná verzia 3.5.
- V prípade systému Windows XP SP1 treba doinštalovať WIA Redistributable 2.0. Knižnicu *wiaaut.dll* je treba skopírovať do *%windir%\System32* a spustiť príkaz *RegSvr32 WIAAut.dll*.

Oba tieto produkty sa spolu so zdrojovým kódom nachádzajú na CD.